

UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE CIENCIAS BIOLÓGICAS



TESIS DOCTORAL

**Filogenia, filogeografía y biogeografía histórica de las
especies mediterráneas del género *Squalius* Bonaparte,
1837 (Actinopterygii, Cyprinidae)**

MEMORIA PARA OPTAR AL GRADO DE DOCTORA

PRESENTADA POR

Silvia Perea Aranda

DIRECTOR

Ignacio Doadrio Villarejo

Madrid, 2018

UNIVERSIDAD COMPLUTENSE DE MADRID

Facultad de Ciencias Biológicas



Filogenia, Filogeografía y Biogeografía Histórica de las especies mediterráneas del género *Squalius* Bonaparte, 1837 (Actinopterygii, Cyprinidae)

Memoria presentada para optar al grado de Doctor por

Silvia Perea Aranda

Bajo la dirección del Doctor

Ignacio Doadrio Villarejo

Madrid – Abril 2017

Ignacio Doadrio Villarejo, Profesor en Investigación del Consejo Superior de Investigaciones Científicas

CERTIFICA

Que la presente memoria titulada Filogenia, Filogeografía y Biogeografía Histórica de las especies mediterráneas del género *Squalius* Bonaparte, 1837 (Actinopterygii, Cyprinidae), que para optar al grado de Doctor presenta Silvia Perea Aranda, ha sido realizada bajo mi dirección en el Departamento de Biodiversidad y Biología Evolutiva del Museo Nacional de Ciencias Naturales – CSIC (Madrid). Esta memoria está además adscrita a la Facultad de Ciencias Biológicas de la Universidad Complutense de Madrid. Considerando que representa trabajo suficiente para constituir una tesis doctoral, autorizamos su presentación.

Y para que así conste firmamos el presente certificado.

El Director:

El doctorando

Ignacio Doadrio Villarejo

Silvia Perea Aranda

En Madrid a 6 de Abril de 2017

El trabajo de esta Tesis Doctoral ha podido llevarse a cabo con la financiación de los proyectos CGL2010-15231BOS del Ministerio de Ciencia e Innovación.

ÍNDICE

RESUMEN	1
ABSTRACT	3
I. INTRODUCCIÓN GENERAL	5
II OBJETIVOS	91
III PUBLICACIONES	
CAPÍTULO I	95
Cenozoic and climatic events in southern Iberian Peninsula: implications for the evolutionary history of freshwater fish of the genus <i>Squalius</i> (Actinopterygii, Cyprinidae)	
CAPÍTULO II:	135
Influence of Iberian paleogeography on the evolutionary history of two Iberian endemic <i>Squalius</i> species (Actinopterygii, Cyprinidae): a multilocus phylogenetic and biogeographical approach.	
CAPÍTULO III:	187
The Dinaric karst is driving the diversification of the Eastern Adriatic <i>Squalius</i> Bonaparte, 1837 (Actinopterygii, Cyprinidae) species belonging to the Mediterranean genetic lineage: evidence from mitochondrial and nuclear markers	
CAPÍTULO IV:	205
Phylogeography, historical demography and hábitat suitability modeling of freshwater fishes inhabiting seasonally fluctuating Mediterranean systems: a case study using the Iberian cyprinid <i>Squalius valentinus</i>	
CAPÍTULO V:	253
Ancient mitochondrial capture as factor promoting mitonuclear discordance in freshwater fishes: a case study in the genus <i>Squalius</i> (Actinopterygii, Cyprinidae) in Greece	
IV DISCUSIÓN GENERAL	293
V CONCLUSIONES	315

RESUMEN

El género *Squalius* Bonaparte 1837, es uno de los géneros más diversificados y ampliamente distribuidos en la región Mediterránea, además, es uno de los géneros de ciprínidos más antiguo de la península ibérica ya que se conocen fósiles desde el Oligoceno. Por tanto, su historia evolutiva debe estar fuertemente ligada a los procesos históricos y climáticos de la región Mediterránea durante la mayor parte del Terciario y Cuaternario. El género *Squalius* presenta dos linajes evolutivos desde un punto de vista genético separados por la cordillera alpina como principal barrera generadora de patrones vicariantes entre ambos linajes. La presente tesis ha tenido como objetivo estudiar los distintos patrones de diversificación de las especies que pertenecen al linaje mediterráneo así como de los procesos que los originan y que han dado lugar a la alta diversidad que aparece actualmente en las penínsulas mediterráneas. De esta forma, se estudian las relaciones filogenéticas entre las distintas especies mediterráneas del género. Nuestros resultados sugieren que, a pesar de los estudios previos sobre estas especies, todavía quedan cuestiones evolutivas de gran interés por resolver, especialmente en la península ibérica, donde la diversidad de linajes evolutivos es más alta de lo que se consideraba previamente. A través del análisis filogenético de estas especies ibéricas del género, esta tesis también muestra que las aproximaciones bayesianas de delimitación de especies, que tienen en consideración la coalescencia de las poblaciones, resuelven las relaciones filogenéticas entre las especies de forma más precisa que otros métodos filogenéticos, aunque estos últimos manifiestan de forma más evidente las incongruencias entre las topologías de los árboles. De esta forma, cuando existen procesos de hibridación y especiación reticulada los árboles bayesianos de especies no pudieron ser generados como ocurrió en el caso de las especies griegas en las que existe una captura mitocondrial. Por otro lado, los resultados de esta tesis muestran un patrón de especiación alopátrica reciente en las especies del género de la vertiente oriental del Adriático guiado por la finalización de la formación del karst Dinárico, indicando la influencia de los procesos geomorfológicos en la historia evolutiva de los organismos vivos. En esta tesis también se evaluó la estructura geográfica y los patrones de diversidad genética de las especies ibéricas del género *Squalius* demostrando la importancia de los factores climáticos en la generación de estos patrones, siendo los dos factores climáticos más relevantes la estacionalidad del clima mediterráneo y las glaciaciones pleistocénicas. Así, se encuentra un patrón de diversidad genética en la península ibérica congruente con estudios previos de forma que las poblaciones del sur presentan mayor diversidad genética que las poblaciones del norte. En las poblaciones del sur pertenecientes a *S. pyrenaicus*, la estacionalidad mediterránea ha favorecido una estructura metapoblacional como consecuencia de continuos cuellos de botella y fragmentación de las poblaciones. En las poblaciones del norte adscritas a *S. carolitertii*, la pérdida de variabilidad genética puede ser una consecuencia de fuertes reducciones poblacionales e incluso extinción local de poblaciones provocada por el efecto de las glaciaciones. Un patrón de diversidad genética diferente se halla en las poblaciones del levante

español de *S. valentinus*, donde la mayor diversidad genética aparece en los tramos bajos de los grandes ríos y en pequeños ríos que presentan manantiales de caudal y temperatura constantes. Estos ambientes proporcionan una mayor estabilidad ambiental para las poblaciones. Por tanto, en este caso, la variabilidad genética parece estar influenciada no solo por factores climáticos sino también por la estructura y geomorfología de los ríos donde habitan estas especies, como también se observa en el sur de la península ibérica donde la laguna de la Janda posiblemente actuó como un reservorio de variabilidad genética de la fauna de peces de agua dulce de esta región.

ABSTRACT

The genus *Squalius* Bonaparte 1837, is one of the most diversified and widely distributed genus of the Mediterranean region. Besides, it is one of the oldest cyprinid genera in the Iberian Peninsula and fossils from the Oligocene period are known. Therefore, its evolutionary history is strongly linked to the Cenozoic historical and climatic processes of the Mediterranean region. The genus *Squalius* is represented by two different evolutionary lineages isolated by the biogeographic barrier that supposed the Alpine range. The main focus of this thesis is studying the species that belong to the Mediterranean lineage as well as the processes that have giving rise to the high current diversity of this lineage in the Mediterranean peninsulas. Thus, the results of the present thesis suggests that in spite of the number of scientific studies focus on these Mediterranean species there are still many interesting evolutionary questions to be resolved, especially in the Iberian Peninsula, where diversity of genetic lineages is higher than previous knowledge. This thesis also shows that Bayesian species delimitation approaches are more powerful for deciphering relationships within species than other phylogenetic methods, although these latter methods are also important to highlight incongruences between tree topologies. Therefore, when hybridization processes or reticulated hybridization were present, species trees could not be generated, as was the case of the mitochondrial capture inferred in the Peloponeso. On the other hand, results from this thesis showed a recent process of allopatric speciation drove by the final episodes of the Dinaric karst configuration, indicating the influence of geomorphological processes in the evolutionary history of organisms. The genetic structure and diversity patterns of the Iberian species of the genus *Squalius* were also inferred; demonstrating the importance of climatic factors in the generation of those processes and being the two most relevant climatic factors are the Mediterranean seasonality and the Pleistocene glaciations. Thus, a genetic diversity pattern congruent with previous studies was found. This pattern showed that southern populations presented higher values of genetic diversity than northern populations. In the southern populations, belonging to the species *S. pyrenaicus*, the Mediterranean seasonality has favoured a metapopulation structure as consequences of continuous bottlenecks and fragmentation of populations. In northern populations belonging to *S. carolitertii*, the loss of genetic variability could be a consequence of drastic population declines or even local extinction of populations caused by the effect of Pleistocene glaciations. A different pattern of genetic diversity is found in the populations of the Spanish Levant region, where the highest genetic diversity was found in populations from the lower reaches of the larger rivers and in small karstic springs with permanent caudal and stable climatic conditions. These environments provide a higher stability for these populations. Therefore, in this case, the genetic variability is also influenced by geomorphological features of the rivers inhabited by these populations. This is also observed in the south of the Iberian Peninsula where the Janda Lagoon probably acted as a reservoir of genetic variability for the freshwater fish fauna of this region.

I. INTRODUCCIÓN GENERAL

1. Evolución biológica e historia evolutiva

Desentrañar la historia evolutiva de los seres vivos actuales es el principal objetivo de la biología evolutiva, cuyo fin es analizar y comprender los procesos que han originado su diversidad actual a partir de un antepasado común, y que siguen transformando dicha diversidad a lo largo del tiempo dando lugar a la aparición de nuevas formas de vida. La biología evolutiva busca además conocer las relaciones de parentesco entre los distintos organismos (filogenia). Sin embargo, esclarecer la historia evolutiva de la mayoría de los grupos taxonómicos actuales, o sus relaciones filogenéticas, es un desafío, ya que la información que se tiene de muchos de ellos es incompleta. Metafóricamente hablando, investigar la historia evolutiva es intentar resolver un rompecabezas en el que faltan la mayoría de las piezas, especialmente aquellas relacionadas con el registro fósil, y en el que se tienen que ir encajando unas con otras de forma coherente a medida que se van adquiriendo al aumentar el conocimiento científico.

Tradicionalmente la evolución biológica se ha estudiado a dos escalas diferentes: microevolución y macroevolución. La microevolución se asocia a los procesos que dan lugar a la especiación y a la extinción de los organismos, mientras que la macroevolución es vista como un proceso histórico que analiza la repetición de patrones evolutivos a un nivel específico o superior. Estos términos fueron acuñados por Filipchenko en la años 20 (1927, 1929) y aunque en los últimos años la delimitación de ambos conceptos comienza a ser cada vez más difusa, ya que las fuerzas evolutivas que actúan sobre la macroevolución (mutación, selección, migración y deriva genética) son las mismas que intervienen en los procesos microevolutivos, permanece la diferencia en la escala temporal que sigue siendo un criterio epistemológico importante (Dobzhansky, 1937). Sin embargo, el debate sobre si macroevolución y microevolución son equivalentes, siendo su única diferencia la escala de tiempo, sigue abierto ya que, de acuerdo a algunos autores, algunas de las propiedades de la macroevolución no pueden ser explicadas simplemente como una acumulación de procesos microevolutivos, sino que existen unas propiedades emergentes asociadas al proceso de macroevolución (Mayr, 1963, 1982; Avise *et al.*, 1987; Jablonski, 2000, 2008; Hendry y Kinnison, 2001; Rabosky y McCune, 2010).

La macroevolución trata sobre los procesos evolutivos que ocurren a una escala evolutiva grande, es decir, aquellos que tienen lugar en grupos taxonómicos jerárquicamente altos (especies, familias, órdenes, etc) y que suponen un cambio evolutivo drástico dentro de los mismos, observable sólo después de que transcurra un largo periodo de tiempo (Dobzhansky, 1937; Rabosky y McCune, 2010). La macroevolución se encarga, por tanto, del estudio de las novedades evolutivas (Shubin *et al.*, 2009; Hallgrímsson *et al.*, 2012); los grandes patrones evolutivos como las tasas de cambio o las radiaciones evolutivas (Harmon *et al.*, 2003;

Seehausen, 2004; Brawand *et al.*, 2014); la formación y extinción de linajes (Mittelbach *et al.*, 2007; Burbrink y Pyron, 2010; Antonelli y Sanmartín, 2011); y la disparidad morfológica entre grupos taxonómicos (Derek *et al.*, 1992; Brusatte *et al.*, 2011). Las interacciones bióticas interespecíficas también deben ser tenidas en cuenta a la hora de considerar los procesos macroevolutivos en su interacción con la disciplina ecológica (Jablonski, 2008).

Por el contrario, los procesos microevolutivos serían aquellos que ocurren en un nivel inferior al de especie, es decir aquellos procesos que tienen lugar entre y dentro de poblaciones, y dan lugar a cambios más pequeños y progresivos que también se acumulan con el paso del tiempo (Templeton, 2006). Los procesos microevolutivos provocan cambios en las frecuencias génicas de una población biológica en una escala de tiempo pequeña (Fig. 1). Las mutaciones causan cambios en los genotipos de una generación que se transmiten a la descendencia y se van acumulando con el paso del tiempo de una forma neutral (Templeton, 2006). Sin embargo, los cambios microevolutivos también están relacionados con procesos de selección, migración y deriva genética (Templeton, 2006). A escala microevolutiva, la selección tiene lugar cuando ocurre una reproducción diferencial entre los individuos de una población biológica de forma que solamente se transmiten a la descendencia unos genotipos determinados, los mejor adaptados a las condiciones ambientales donde se desarrollan, aumentando la frecuencia de dichos genotipos en la población y en último término su fijación si el resto de genotipos “peor adaptados” desaparece de la misma (selección positiva; Templeton, 2006). Por el contrario, la selección puede actuar en contra de un determinado alelo, que irá disminuyendo su frecuencia a medida vayan pasando las generaciones (selección negativa; Templeton, 2006). Un tercer ejemplo de selección se da cuando los individuos heterocigotos presentan ventajas adaptativas respecto a los heterocigotos (selección balanceada; Templeton, 2006). La deriva genética es otro proceso que a escala microevolutiva produce fijación de genotipos a lo largo del tiempo, pero en este caso como consecuencia del azar (Templeton, 2006). Un claro ejemplo de deriva genética es el que tiene lugar cuando ocurren fenómenos de cuello de botella, en los que tras una drástica reducción de la población cuando ésta se recupera la frecuencia génica de la población ha cambiado totalmente respecto a la población original (Nei *et al.*, 1975). Por último, la migración provoca el intercambio de genotipos entre poblaciones distintas dentro de una misma especie como consecuencia del movimiento de individuos de unas poblaciones a otras, cambiando por tanto las frecuencias génicas de una generación a la siguiente (Allendorf, 1983).

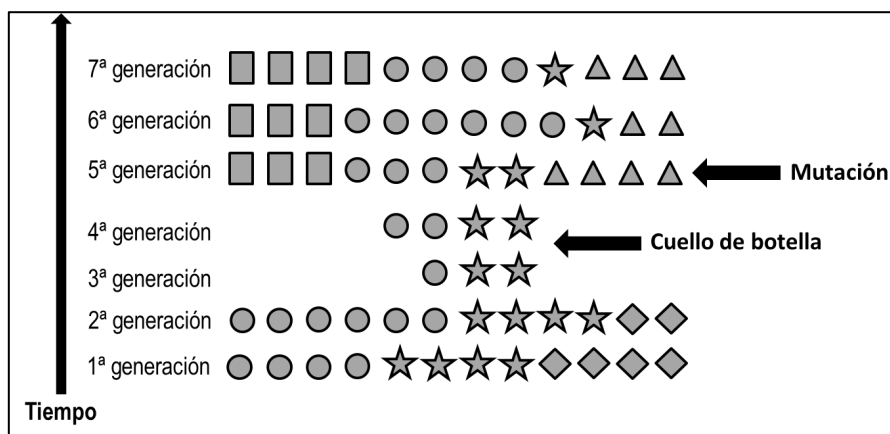


Fig. 1. Representación esquemática d cambios en las frecuencias alélicas a lo largo del tiempo como consecuencia de procesos microevolutivos. Cada símbolo representa un alelo diferente. En la tercera generación se representa un cuello de botella que muestra una drástica reducción poblacional y la pérdida de uno de los alelos (rombo). En la quinta generación se representa la aparición de un nuevo alelo (triángulo), consecuencia de un proceso de mutación.

2. La familia Cyprinidae

2.1. Biología y Ecología de la familia Cyprinidae

La familia Cyprinidae se encuentra ampliamente distribuida por toda África, la región Indo-Malaya y el Paleártico y está ausente en Sudamérica, Madagascar, Nueva Zelanda y Australia. Es una de las familia más diversificadas de peces de agua dulce junto a los Gobiidae, y es la que presenta el mayor número de especies dentro de los vertebrados (aproximadamente 3000; Eschmeyer *et al.*, 2016; Nelson *et al.*, 2016). La diversidad específica de los ciprínidos es más alta en Eurasia, donde se cita aproximadamente el 80% de las especies (Gaubert *et al.*, 2009; Nelson *et al.*, 2016).

Los ciprínidos abarcan un rango muy amplio de tallas corporales, en el que destaca el pequeño tamaño de la especie *Paedocypris progenetica* Kottelat, Britz, Tan y Witte, 2006, con aproximadamente 10 mm de longitud estándar (Kottelat *et al.*, 2006) o el gran tamaño de la especie *Catlocarpio siamensis* Boulanger, 1898, que puede sobrepasar los tres metros de longitud estándar (Baird *et al.*, 1999). Respecto a la forma del cuerpo, los ciprínidos pueden dividirse en dos morfotipos generales, algunas especies presentan un cuerpo alto, comprimido lateralmente, mientras que otras especies tienen una forma fusiforme y más alargada; sin embargo, es posible encontrar una gran disparidad morfológica dentro de los ciprínidos. Así, por ejemplo, en la familia Danionae se ha observado una tendencia hacia la miniaturización y reducción de caracteres morfológicos en varios de sus géneros como *Danionella* Roberts, 1986,

Paedocypris Kottelat, Britz, Tan y Witte, 2006 y *Sundadanio* Kottelat y Witte, 1999 (Kottelat *et al.*, 2006; Britz *et al.*, 2011; Britz y Conway, 2015). Otros géneros como *Garra* Hamilton, 1822 o *Labeo* Cuvier, 1816 presentan las estructuras bucales modificadas en forma de un órgano succionador (Nelson *et al.*, 2016).

Como miembros del superorden Ostariophysi, los ciprínidos se caracterizan por la presencia de una estructura anatómica denominada aparato de Webber, que conecta la vejiga natatoria con el sistema auditivo (Winfield y Nelson, 1991). El rasgo morfológico más característico que define a la familia Cyprinidae es la presencia de dientes faríngeos, ya que carecen de dientes en la mandíbula y en su lugar han desarrollado una compleja estructura asociada a la alimentación y a la función masticatoria en el quinto arco branquial (Fig. 2), constituida por un hueso faríngeo bien desarrollado en forma de una placa dentaria sobre la que se asientan dientes, en número variable y dispuestos en una, dos o hasta tres filas (Chu, 1935; Winfield y Nelson, 1991; Cavender y Coburn, 1992). La morfología del hueso faríngeo y de sus dientes es variable en función del tipo de alimentación de cada especie (Eastman, 1977). El número de dientes y su disposición sobre la placa dentaria se denomina fórmula dentaria y es relativamente constante a nivel de especie. La forma del hueso faríngeo y sus dientes son las piezas esqueléticas que han sido más utilizadas en taxonomía y sistemática de ciprínidos (Chu, 1935; Lozano Rey, 1935; Maitland y Linsell, 1980; Cavender, 1991; Cavender y Coburn, 1992). Otras características morfológicas que definen a la familia Cyprinidae son las siguientes: mandíbula superior protrusible con un hueso medio entre las mandíbulas llamado kinetmoides, basioccipital con un largo proceso posterior que encierra la aorta dorsal, un muy desarrollado 5º músculo elevador que se origina en una profunda fosa subtemporal, a menudo barbillas bucales presentes, algunas especies presentan radios espinosos en la aleta dorsal y siempre carecen de escamas en la región cefálica (Nelson *et al.*, 2016). Sin embargo no se han identificado sinapomorfías exclusivas de la familia Cyprinidae (Howes, 1991).

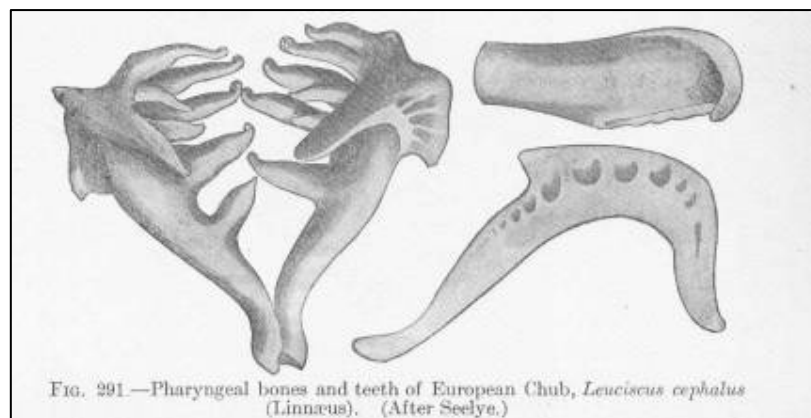


Fig. 4 Dientes faríngeos de *Squalius cephalus* (Fuente: Freshwater and Marine Image Bank at the University of Washington; Licencia: dominio público).

El espectro ecológico de los ciprínidos también es muy amplio, ocupando todo tipo de ecosistemas de agua dulce, desde hábitats lénticos como lagos o lagunas y áreas inundadas de zonas costeras a ambientes lóticos como ríos y torrentes de alta montaña, así como toda la gama intermedia (Winfield y Nelson, 1991; Nelson *et al.*, 2016). Respecto a su alimentación, esta es muy variada, muchos géneros son insectívoros, otros son fitófagos y la familia también tiene representantes omnívoros; algunos constituyen grupos tróficos más especializados, como ocurre, entre otros géneros, con *Capoeta* Valenciennes, 1842, *Varicorhinus* Rüppell, 1835 o *Cyprinion* Heckel, 1843, que se alimentan raspando las algas adheridas a la superficie de las rocas o del sustrato, teniendo estructuras bucales adaptadas a tal fin (Berg, 1949; Groenewald, 1958). La gran mayoría de ciprínidos son peces primarios, es decir, están restringidos al medio dulceacuícola ya que no poseen mecanismos de osmorregulación que les permitan colonizar el medio marino (Myers, 1938); pero algunos géneros como *Tribolodon* Sauvage, 1883 tienen una amplia capacidad de osmorregulación pudiendo habitar y desplazarse a través de zonas de estuario e incluso marinas (Dibowsky, 1872; Nikolskii, 1889; Sasaki *et al.*, 2007; Imoto *et al.*, 2013). Respecto a sus modos de reproducción, los ciprínidos son ovíparos, presentan fecundación externa y la mayoría de especies no cuidan de las puestas, aunque los machos de algunas especies construyen nidos y protegen los huevos (Winfield y Nelson, 1991; Pyron, 1996; Nelson *et al.*, 2016). Al contrario que ocurre con otras familias de peces de agua dulce como Cyprinodontidae o Goodeidae, el dimorfismo sexual no es muy acusado en los ciprinidos y solo se aprecia en rasgos puntuales generalmente asociados a la longitud de las aletas o coloración de las mismas durante la época reproductiva, aunque algunos géneros como *Paedocypris* muestran un dimorfismo muy marcado (Douglas, 1993; Pyron, 1996; Kottelat *et al.*, 2006; Gante *et al.*, 2007). En muchas especies los machos desarrollan tubérculos nupciales sobre el cuerpo, que generalmente son más numerosos en la cabeza (Koehn, 1965; Chen, 1996). La alta fecundidad de la mayoría de las especies da origen a la etimología de la familia que proviene de la diosa Cipris, también llamada Afrodita.

Otra característica de los ciprínidos es la presencia de diferentes ploidías, con una especie *Ptychobarbus dipogon* (Regan, 1905) dentro de los vertebrados con el número cromosómico más alto ($2n=446$) y siendo frecuentes los géneros con especies diploides ($2n=50$), triploides ($2n=75$), tetraploides ($2n=100$), y hexaploides ($2n=150$), como es el caso de los géneros *Barbus* Daudier, 1805, *Capoeta* Valenciennes, 1842, *Carasobarbus* Karaman, 1971, *Cyprinus* Linnaeus, 1758 y *Luciobarbus* Heckel, 1843 (Ohno *et al.*, 1967; Wolf *et al.*, 1969; Collares-Pereira, 1994; Machordom y Doadrio, 2001a). La poliploidización en la familia Cyprinidae, al igual que en otras familias de peces en las que se ha producido este mecanismo evolutivo, es de origen aloploidice, es decir procede de la duplicación de genomas por

hibridación de dos o más especies (Comber y Smith, 2004); aunque en la mayoría de los casos se desconoce la procedencia de los genomas parentales, pudiendo incluso pertenecer estos a una especie actualmente extinta, como ocurre en el género *Squalius* (Carmona *et al.*, 1997; Alves *et al.*, 2001). La duplicación de genomas ha supuesto un trascendental paso evolutivo para muchos organismos vivos, ya que incrementa la fuente de material genético sobre la que actúan procesos como la mutación, la deriva genética o la selección, brindando a los organismos poliploides nuevas oportunidades evolutivas así como el potencial para generar variación genética adaptativa, incorporar novedades evolutivas e incluso crear nuevas especies (Lewontin y Birch, 1966; Arnold, 1997, 2006; Crow y Wagner, 2006). En el caso de los peces, la poliploidización se ha asociado a un mayor tamaño corporal, una tasa de crecimiento más rápida, una vida más longeva y una mayor adaptabilidad a nuevas condiciones ambientales (Uyeno y Smith, 1972; Schulz, 1980). A un nivel molecular más reduccionista, la poliploidización también se ha relacionado con alteraciones en los reordenamientos genómicos que a su vez puede tener efecto en los tipos de cambios evolutivos que se producen y en la tasa a la que se producen (Comber y Smith, 2004). Así mismo, a una escala macroevolutiva, la poliploidización ha sido considerada un factor promotor de grandes radiaciones evolutivas, ya que los órdenes de peces que tienen una mayor riqueza de especies son aquellos que presentan formas poliploides (Comber y Smith, 2004). Uno de los ejemplos más complejos de ploidías se encuentra en el género *Squalius* Bonaparte, 1843, objeto de la presente tesis doctoral, en el que una de sus especies, *S. alburnoides* Steindachner, 1866, presenta poblaciones diploides, triploides y tetraploides, asociadas a su vez a distintos morfotipos y a complejas estrategias reproductivas, incluyendo la reproducción asexual (Carmona *et al.*, 1997; Cunha *et al.*, 2004, 2009; Sousa-Santos *et al.*, 2006). Sin embargo, por su historia evolutiva particular, *S. alburnoides* no se ha incluido en esta tesis y todas las especies de *Squalius* que se han tenido en cuenta son diploides (2N=50).

Desde un punto de vista socioeconómico, los ciprínidos han suscitado siempre un gran interés. De esta forma, la subfamilia Cyprininae, a la que pertenece la carpa (*Cyprinus carpio* L., 1758) o el pez rojo (*Carassius auratus* L., 1758), incluye muchas especies utilizadas en pesquerías, acuicultura y en acuariofilia, (Nelson *et al.*, 2016). Esto ha conllevado a un comercio en continua expansión de estas especies que ha llevado asociado un creciente trasiego de las mismas a nivel global, que a su vez ha derivado en que algunas especies de ciprínidos, como es el caso de la carpa, estén incluidas entre las cien especies invasoras más dañinas (Lowe *et al.*, 2010). Además, un ciprínido, *Danio rerio* (Hamilton, 1822), es utilizado como especie modelo en multitud de disciplinas científicas (ver por ejemplo Briggs, 2002; Grunwald y Elsen, 2002; Norton y Bally-Cuif, 2010).

2.2. Filogenia y Sistemática de la Familia Cyprinidae

Aunque la monofilia de la familia Cyprinidae está apoyada por filogenias mitocondriales y nucleares robustas (Liu *et al.*, 2002; Liu y Chen, 2003; Saitoh *et al.*, 2006, 2011; Wang *et al.*, 2007; He *et al.*, 2008; Chen y Mayden, 2009; Mayden y Chen, 2010), las relaciones filogenéticas dentro de la familia Cyprinidae todavía no están completamente resueltas. La clasificación de los ciprínidos ha sido constantemente debatida y se han reconocido de dos a doce subfamilias dependiendo del autor y de los caracteres morfológicos analizados (Bonaparte, 1846; Nikolsky, 1954; Gosline, 1978; Howes, 1980, 1983; Arai, 1982; Chen *et al.*, 1984; Cavender y Coburn, 1992; Arai y Kato, 2003). Sin embargo, estas clasificaciones morfológicas muchas veces entran en conflicto cuando se trata de integrarlas en clasificaciones y filogenias obtenidas desde una perspectiva molecular (Briolay *et al.*, 1997; Zardoya y Doadrio, 1999; Gilles *et al.*, 2001; Cunha *et al.*, 2002; Liu *et al.*, 2002; Liu y Chen, 2003; Saitoh *et al.*, 2006, 2011; He *et al.*, 2008; Chen y Mayden, 2009; Bufalino y Mayden, 2009, 2010a,b; Rüber *et al.*, 2007; Gaubert *et al.*, 2009; Zhang *et al.*, 2009; Mayden y Chen, 2010; Perea *et al.*, 2010; Levin *et al.*, 2012; Wang *et al.*, 2012; Yang *et al.*, 2015), ya que muchos de los caracteres comúnmente utilizados en los estudios morfológicos de ciprínidos, como las barbillas o dientes faríngeos, son susceptibles de sufrir homoplasias y generalmente no está clara la homología entre algunos de estos caracteres morfológicos (Bogutskaya, 1990a,b; Howes, 1991; Zardoya y Doadrio, 1999). Clasificaciones posteriores basadas en caracteres moleculares incluso han propuesto la elevación de estas subfamilias al rango de familias (Chen y Mayden, 2009; Mayden y Chen, 2010). Sin embargo, después de las filogenias nucleares basadas en alozimas (Doadrio *et al.*, 1998; Doadrio y Carmona, 1998; Machordom *et al.*, 1995, 1998; Ketmaier *et al.*, 1998, 2003; Tsigenopoulos *et al.*, 1999; Hänfling y Brandl, 2000; Ito *et al.*, 2002; Kartatsev *et al.*, 2002; Sakai *et al.*, 2002, 2006; Kartavtsev y Hanzawa, 2007), los estudios que analizan secuencias de ADN nuclear no son muy numerosos y no incluyen a todos los grupos debido a la poliploidía de algunos de los géneros de ciprínidos como *Barbus*, *Luciobarbus* o *Capoeta* (He *et al.*, 2008; Chen y Mayden, 2009; Bufalino y Mayden, 2010a,b; Mayden y Chen, 2010; Perea *et al.*, 2010; Wang *et al.*, 2012; Yang *et al.*, 2015), por lo que la mayoría de las hipótesis filogenéticas formuladas para estos géneros poliploides están basadas exclusivamente en ADN mitocondrial, de manera que no pueden extraerse conclusiones taxonómicas y sistemáticas definitivas para todos los grupos (Machordom y Doadrio, 2001a,b; Levin *et al.*, 2012; Ghanavi *et al.*, 2016). Hay que tener en cuenta que la taxonomía y sistemática de la familia Cyprinidae ha sido y es en general muy dinámica y cambiante debido a que sólo se han estudiado desde el punto de vista filogenético un porcentaje muy pequeño de todas las especies que se estima que componen la familia (Nelson *et al.*, 2016, Eschmeyer *et al.*, 2016).

Una de las subfamilias de ciprínidos es la subfamilia Leuciscinae (Familia Leuciscidae de acuerdo a Chen y Mayden, 2009), a la que pertenece el grupo taxonómico objeto de estudio de esta tesis, el género *Squalius* Bonaparte, 1837. Los caracteres diagnósticos de la subfamilia son los siguientes: ausencia de barbillas; dientes faríngeos generalmente colocados en una o dos filas; aleta dorsal corta en la que el último radio simple nunca está aserrado ni es espinoso; y además presentan otra serie de características osteológicas que los diferencian del resto de subfamilias (Howes, 1991; Kottelat y Freyhof, 2007a). La monofilia de la subfamilia Leuciscinae ha sido tradicionalmente cuestionada (Chen *et al.*, 1984; Howes, 1991; Cavender y Coburn, 1992; He *et al.*, 2008; Nelson *et al.*, 2016). Muchas especies de leuciscinos son morfológicamente similares, incluso especies pertenecientes a distintos géneros (Howes, 1991), por lo que muchas de ellas han sido transfiriendas intermitentemente de unos géneros a otros, tal como por ejemplo ocurre con distintas especies originalmente descritas o en algún momento consideradas como pertenecientes al género *Rutilus* Rafinesque, 1820 (Kottelat, 1997; Ketmaier *et al.*, 1998; Doadrio y Carmona, 2003; Bogutskaya *et al.*, 2006; Freyhof *et al.*, 2006; Kottelat y Freyhof, 2007a; Freyhof y Özuluğ, 2010; Bianco y Ketmaier, 2014). Igual sucede con el género *Pseudophoxinus* Bleeker, 1860, el cual ha demostrado ser polifilético (Perea *et al.*, 2010), de hecho, muchas de las especies previamente adscritas a este género (Collares-Pereira y Almaça, 1979; Coad, 1981; Economidis, 1991, 1995; Kottelat y Barbieri, 2004), actualmente han pasado a formar parte de otros géneros como *Acanthobrama* Heckel, 1843, *Anaocypris* Collares-Pereira, 1983, *Pelagius* Kottelat y Freyhof, 2007, *Telestes* Bonaparte, 1840 o *Tropidophoxinellus* Stephanidis, 1974 (Collares-Pereira, 1983; Zardoya y Doadrio, 1999; Kottelat y Freyhof, 2007a,b; Perea *et al.*, 2010). El propio género *Squalius* fue reconocido durante mucho tiempo como un subgénero del género *Leuciscus* (Berg, 1949; Banareescu, 1964), que posteriormente fue elevado al rango de género (Doadrio, 2001; Sanjur *et al.*, 2003). Más compleja y confusa, y fuera del marco teórico de esta tesis, es todavía la clasificación morfológica de las tribus dentro de la subfamilia Leuciscinae (Travers, 1989; Bogutskaya, 1990a,b; Howes, 1991; He *et al.*, 2008).

Desde un punto de vista molecular, filogenias de mitogenomas completos apoyan la monofilia de la subfamilia Leuciscinae incluyendo los representantes eurasiáticos junto con la especie norteamericana *Notemigonus crysoleucas* (Saitoh *et al.*, 2006, 2011). Mientras que estos autores defienden la inclusión del resto de especies norteamericanas en su propia subfamilia, Phoxininae, otros postulan que la subfamilia Leuciscinae está compuesta por todas las especies euroasiáticas más las norteamericanas (leuciscinos + phoxininos) aunque ambos grupos sean monofiléticos y considerados dos tribus independientes (Pseudaspinini y Leuciscini) dentro de la subfamilia Leuciscinae (Cunha *et al.*, 2002; Kottelat y Freyhof, 2007a; Wang *et al.*, 2007; Imoto *et al.*, 2013). Algunos estudios teniendo en cuenta secuencias de ADN mitocondriales y nucleares aportan desde un punto de vista molecular la inclusión de tres especies africanas circunscritas a

los géneros *Pseudophoxinus* y *Tropidophoxinellus* en la subfamilia Leuciscinae s.s., es decir, excluyendo los phoxininos (Perea *et al.*, 2010).

La posición filogenética de los grupos basales de la subfamilia Leuciscinae s./ (Leuciscinae + Phoxininae) también es controvertida y varía dependiendo de los taxa incluidos, los caracteres morfológicos o moleculares utilizados y el tipo de análisis realizado (Cavender y Coburn, 1992; Cunha *et al.*, 2002; Rüber *et al.*, 2007; Saitoh *et al.*, 2006; Sasaki *et al.*, 2007; Wang *et al.*, 2007; Strange y Mayden, 2009; Zhang *et al.*, 2009; Bufalino y Mayden, 2010a,b; Imoto *et al.*, 2013). La sistemática de la subfamilia a nivel de género tampoco está bien resuelta, ya que no existe congruencia entre los datos morfológicos y moleculares para definir muchos de los géneros, entre ellos *Squalius* (Kottelat y Freyhof, 2007a).

2.3. Patrones biogeográficos generales de la Familia Cyprinidae y la subfamilia Leuciscinae en la Región Mediterránea

El marco biogeográfico de estudio en el que se engloba esta tesis es la región Mediterránea, donde se distribuyen las especies del género *Squalius* que pertenecen al linje evolutivo Mediterráneo de acuerdo a Sanjur *et al.*, 2003. La región Mediterránea es una de las zonas más complejas del mundo desde un punto de vista geológico, geográfico y naturalístico y, además, ha sido reconocida como un punto caliente de biodiversidad con una gran diversidad de especies en todos los grupos taxonómicos (Médail y Quézel, 1999; Myers *et al.*, 2000). Respecto a su origen, este se remonta al mar de Tethys en la era Mesozoica (hace aproximadamente 200 millones de años), y desde entonces, diversos episodios geológicos asociados a movimientos tectónicos han causado drásticas modificaciones en esta región hasta alcanzar la configuración geográfica actual de la región Mediterránea (Blondel y Aronson, 1999; Rögl, 1999). La compleja paleogeografía de la región Mediterránea, con arcos de islas migratorios, cinturones tectónicos fragmentándose y otros eventos también relacionados con la tectónica de placas ha contribuido a la formación de patrones biogeográficos reticulados y complejos en los cuales los actuales grupos de peces de agua dulce, entre ellos el género *Squalius*, han ido evolucionando a medida que diferentes barreras biogeográficas aparecían y desaparecían a lo largo del tiempo evolutivo, aislando unos linajes de otros. De hecho, las regiones Mediterránea y de Asia occidental, donde se distribuye el género completo, han sido consideradas importantes áreas para la diversificación de fauna continental durante el gradual cierre del mar de Tethys (Berg, 1949; Coad, 1987, 1996; Zardoya y Doadrio, 1998, 1999; Durand *et al.*, 1999, 2000, 2002; Tsigenopoulos y Berrebi, 2000; Doadrio y Carmona, 2003, 2004; Perea *et al.*, 2010).

Los peces de agua dulce primarios (Myers, 1938) constituyen un modelo idóneo para estudiar patrones biogeográficos históricos debido a su intolerancia o baja tolerancia a la salinidad, que hace que queden confinados a las cuencas hidrográficas que habitan y por tanto su historia evolutiva va estar condicionada por la historia paleogeográfica del área en la que residen, constituyendo un excelente modelo para estudiar como interaccionan procesos recientes y pasados en la estructura de las poblaciones (ver Bianco y Nordlie, 2008 para una crítica a la división tradicional entre peces primarios y secundarios). De esta forma, ha sido posible testar hipótesis vicariantes analizando sus relaciones filogenéticas o filogeográficas (Rosen, 1975, 1978; Mayden, 1988; Banareescu, 1990; Doadrio, 1990; 1994; Doadrio y Carmona, 1998, 2003; Zardoya y Doadrio, 1998, 1999; Avise, 2000; Šlechtová *et al.*, 2004; Böhlen *et al.*, 2006; Burrridge *et al.*, 2006, 2007; Filipe *et al.*, 2009; Levy *et al.*, 2009; Perea *et al.*, 2010; Imoto *et al.*, 2013).

Los eventos de dispersión son más difíciles de testar formalmente mediante contraste de hipótesis pero pueden inferirse, con un apoyo estadístico asociado, a través de las metodologías de reconstrucción de áreas ancestrales (Near y Keck, 2005; Chen *et al.*, 2013). Al igual que la vicarianza es un factor determinante en procesos de divergencia alopátrica entre poblaciones, que da lugar a especiación cuando transcurre el suficiente tiempo evolutivo (Ronquist, 1997; Lossos y Glor, 2003), la dispersión, bien sea entre sistemas fluviales dulceacuícolas (Banareescu, 1990) o a través del medio marino en aquellas especies tolerantes a la salinidad o que presenten parte de su ciclo vital en aguas marinas como las especies diadromas (Waters *et al.*, 2000), también ha jugado un papel muy importante explicando patrones biogeográficos en peces de agua dulce que no se ajustan a procesos vicariantes como colonización de nuevas áreas o contactos secundarios entre cuencas durante las glaciaciones pleistocénicas (Zemlak *et al.*, 2008; Hughes *et al.*, 2009).

El registro fósil de los ciprínidos más antiguo data del Eoceno (55-33 Ma), con restos que pueden ser adscritos a Cyprininae, Barbinae, Labeoninae y Schizothorazinae, y por tanto, se han considerado a los ciprínidos originarios del este y sureste asiático (Zheng, 1962; Thang *et al.*, 1985; Sytchevskaya, 1986; Cavender, 1991; Winfield y Nelson, 1991; Gaudant, 2002; Schulz-Mirbach y Reichenbacher, 2006; Chang y Chen, 2008), donde además se encuentra la mayor diversidad de géneros y especies (Banareescu y Coad, 1991; Nelson *et al.*, 2016; Rainboth, 1991; Gaubert *et al.*, 2009). Los fósiles más antiguos hallados en Europa y Norte América datan del Oligoceno y pertenecen a la subfamilia Leuciscinae (Orbhelová, 1969; Cabrera y Gaudant, 1985; Cavender, 1991; de la Peña, 1995). Las hipótesis biogeográficas clásicas establecidas no pueden explicar por sí solas los patrones biogeográficos tan complejos y reticulados hallados tanto en la familia Cyprinidae como particularmente en la subfamilia Leuciscinae, o como en

algunos de sus géneros analizados de forma independiente, como es el caso del género *Squalius* en el que se centra la presente tesis (Zardoya y Doadrio, 1999; Levy *et al.*, 2009; Perea *et al.*, 2010; Imoto *et al.*, 2013).

Cronológicamente hablando, una de las hipótesis clásicas biogeográficas generales que han sido postuladas para los ciprínidos, denominada “dispersión por el norte” por algunos autores (Doadrio y Carmona, 2003), propone que este grupo de peces colonizó Europa desde Asia por primera vez en el Oligoceno cuando se cerró el Estrecho de Turgai y desapareció el mar de Ob coincidiendo con el levantamiento de los montes Urales, de manera que postula varias olas de colonización por el norte de Europa desde el Oligoceno hasta el final del Plioceno-comienzo del Pleistoceno (35-1,7 Ma) (Banarescu, 1973; 1989; Briggs, 1995). Posteriores capturas fluviales habrían supuesto la colonización de las penínsulas mediterráneas desde el centro de Europa a medida que se fue formando la red hidrográfica europea (Banarescu, 1989; 1992; Economidis y Banarescu, 1991).

Un modelo biogeográfico alternativo denominado “Lago Mare” o “fase lacustre del Mediterráneo”, también dispersalista, plantea que la diversificación de los grupos de la región Mediterránea tuvo lugar al final del Messiniense (5,5-5,3 Ma) durante su fase dulceacuícola tras la desecación del Mediterráneo, cuando este se rellenó de agua procedente del Paratethys, de forma que la ictiofauna continental habría colonizado los márgenes del Mediterráneo durante este periodo (Hsü *et al.*, 1977; Bianco, 1990; Krijgsman *et al.*, 1999). La posterior apertura del Estrecho de Gibraltar en el límite entre el Mioceno y el Plioceno rellenó otra vez el Mar Mediterráneo con agua marina hace 5,3 Ma (Krijgsman *et al.*, 1999; Blanc, 2002; pero ver Carnevale *et al.*, 2006), de manera que de acuerdo a los autores defensores del modelo “Lago Mare”, las nuevas poblaciones establecidas en los márgenes del Mediterráneo quedarían aisladas como consecuencia de fuertes eventos de vicarianza dando lugar a fenómenos de especiación (Penzo *et al.*, 1998; Durand *et al.*, 2003; Tsigenopoulos *et al.*, 2003; Hyuse *et al.*, 2004; Ketmaier *et al.*, 2004).

Hasta el momento la hipótesis biogeográfica de dispersión por el norte no ha podido ser refutada para todos los grupos de ciprínidos, aunque algunos autores consideran este modelo biogeográfico no válido para la subfamilia Leuciscinae (Imoto *et al.*, 2013) y los grupos más antiguos de ciprínidos en la región más occidental de Europa, la Península Ibérica, datan del Oligoceno Tardío y Mioceno Temprano (Cabrera y Gaudant, 1985; de la Peña, 1995), época cercana al cierre del Estrecho de Turgai, aunque algunos autores hipotetizan una conexión antigua entre la Península Ibérica y el norte de Europa como ruta de colonización de la península de los ancestros ibéricos de la subfamilia Leuciscinae (Levy *et al.*, 2009). Sin embargo,

numerosas evidencias sí han refutado la hipótesis dispersalista de Lago Mare demostrando que este modelo no es realista para muchos grupos de agua dulce ya que la diversificación de estos grupos, entre ellos los ciprínidos, es anterior a la fase lacustre del Mediterráneo (Zardoya y Doadrio, 1998, 1999; Doadrio y Carmona, 2003, 2004; Sanjur *et al.*, 2003; Heller *et al.*, 2007; Robalo *et al.*, 2007a; Levy *et al.*, 2009; Perea *et al.*, 2010; Perdices *et al.*, 2016). Esto pone de manifiesto que los patrones biogeográficos que han evolucionado en la fauna de agua dulce de la región Mediterránea no están todavía claros.

La presencia de ciprínidos en el norte de África, donde los fósiles más antiguos hallados datan del Mioceno Medio-Tardío (Baker *et al.*, 1971; Van Couvering, 1977), ha sido explicada a través de conexiones/desconexiones con la Península Ibérica y Asia a lo largo del Cenozoico (Doadrio, 1990; Machordom y Doadrio, 2001b; Perea *et al.*, 2010). De esta forma, conexiones antiguas entre la cordillera Bética en el sur ibérico y las Kabilias en el norte de África han sido postuladas para explicar la distribución actual y las relaciones filogenéticas de las especies de *Luciobarbus* actuales de estas dos regiones biogeográficas (Doadrio y Machordom, 2001b; Doadrio *et al.*, 2016). Por otro lado, relaciones biogeográficas entre la región tropical de Asia y el suroeste asiático también han sido propuestas para algunas de las subfamilias de ciprínidos (Danioninae), siendo la región Mesopotámica (Tigris y Éufrates) el posible punto de conexión histórica entre ambas regiones disjuntas separadas sólo por posibles desertificaciones recientes ocurridas en la región de la península arábiga (Liao *et al.*, 2011). Así mismo, la presencia de especies de *Tropidophoxinellus* en el Peloponeso griego y en el norte de África ha sido explicada por la migración de ancestros comunes al norte de África a través del puente *Gomphoterium*, establecido hace aproximadamente 19 millones de años entre el sureste asiático y África, y posterior vicarianza (Perea *et al.*, 2010). Este puente separó los mares Tethys y Paratehtys y supuso un importante intercambio faunístico de mamíferos entre ambos continentes y, además, ha sido considerado como una entrada a África por grupos asiáticos de ciprínidos relacionados con el grupo Barbinae (Rögl, 1999; Böhme, 2000; Harzhauser *et al.*, 2002, 2007). La presencia de otros ciprínidos emparentados con Alburninae en África, como es el caso de la especie *Pseudophoxinus punicus* (Pellegrin, 1920) todavía no ha sido esclarecida (Perea *et al.*, 2010).

Centrándonos en la subfamilia Leuciscinae *sensu stricto* (a la que pertenece el género *Squalius*) es decir, considerando a Phoxininae una subfamilia independiente y grupo hermano de Leuciscinae (Saitoh *et al.*, 2006, 2011), en base a datos de secuencias de ADN mitocondrial y nuclear se ha demostrado que los grupos más antiguos de leuciscinos divergieron de los phoxininos en el Oligoceno Temprano-Medio, hace aproximadamente 28-33 millones de años (Perea *et al.*, 2010). Esta divergencia de linajes antiguos de leuciscinos, que han originado los

actuales géneros *Delminichthys* Freyhof, Lieckfeldt, Bogutskaya, Pitra y Ludwig, 2006, *Pachychilon* Steindachner, 1882, *Pelagus* o "*Phoxinellus*" *egridiri* Karaman, 1972, se explica como consecuencia de la emergencia de una masa de tierra que ocupaba lo que hoy día es el territorio constituido por la Península Balcánica, Anatolia e Irán, que quedó aislada del resto del continente euroasiático (Rögl, 1999, 2001) y está apoyada por el registro fósil en Anatolia (Howes, 1980; Obrhelová, 1990; Böhme, 1993, 1997; Rögl, 1999, 2001; Böhme e Ilg, 2003). La aparición de estas tierras emergidas y posterior aislamiento se debió a movimientos tectónicos de colisión de placas a lo largo del cinturón alpino e himalayo y dio lugar también a la separación del mar de Tethys del mar epicontinental Paratethys, que se originó en este momento y podría ser congruente con un origen euroasiático de la subfamilia (Cavender, 1991).

Por el contrario, otros autores basados en evidencia obtenida secuenciando mitogenomas completos de especies pertenecientes a las subfamilias Leuciscinae y Phoxininae (Leuciscinae s.l.), sin tener en cuenta las relaciones a nivel nuclear de estos grupos, proponen un origen más antiguo (Cretácico Tardío-Paleoceno; hace aproximadamente 83,3 Ma) y europeo de los leuciscinos, mientras que la aparición de los phoxininos sería posterior (Eoceno; hace aproximadamente 70,7 Ma) una vez que los leuciscinos europeos colonizaron Norte América a través de conexiones transatlánticas (Imoto *et al.*, 2013). Para estas conexiones entre Europa y Norte América se han postulado hipotéticos puentes de tierra favorecidos por descensos del nivel del mar que se mantuvieron emergidos hasta el Eoceno Temprano, hace aproximadamente 55,8 Ma (McKenna, 1983; Tiffney, 1985; Brikiatis, 2014). Uno de estos hipotéticos puentes, denominado *Thulean*, unió el sur de Europa con el este de Norte América a través de las islas Británicas, Groenlandia y otras islas pequeñas próximas al continente americano y ha sido considerado de vital importancia para explicar el intercambio de fauna entre Europa y Norte América al final del Paleoceno entre 57-56 Ma (McKenna, 1983; Tiffney, 1985). Una hipotética conexión terrestre transatlántica más norteña al puente *Thulean* fue el puente De Geer, que favoreció el intercambio faunístico al final del Cretácico y comienzo del Paleoceno (70-65 Ma). Este puente pudo conectar Escandinavia con el este de Norte América y, aunque considerado de menor importancia respecto al puente *Thulean*, también pudo ser fuente de intercambio de fauna entre ambos continentes (Sanmartín, 2001; Brikiatis, 2014). Por tanto, para algunos autores estos dos puentes de tierra han sido considerados como las probables rutas de colonización de los ancestros de los phoxininos desde Europa hacia Norte América rechazando para este grupo taxonómico su llegada de Asia a través de Beringia, ya que de acuerdo a sus tiempos de divergencia estimados, en ese momento Asia y Europa estaban separadas por el Estrecho de Turgai (Imoto *et al.*, 2013).

Interesante es la posición filogenética de la especie *Notemigonus crysoleucas* (Mitchill, 1834), única especie norteamericana de la subfamilia Leuciscinae s.l. estrechamente emparentada con las especies euroasiáticas (Cavender y Coburn, 1992; Cunha *et al.*, 2002; Bufalino y Mayden, 2010a,b; Perea *et al.*, 2010; Imoto *et al.*, 2013). Conexiones transatlánticas apoyadas por la presencia del género fósil *Paleorutilus* Gaudant, 1988, emparentado con los ploxininos norteamericanos, han sido propuestas para explicar su relación con las especies euroasiáticas y no con las norteamericanas (Böhme, 2000; Böhme e Ilg, 2003; Imoto *et al.*, 2013). Se ha estimado que la divergencia del linaje ancestral de *N. crysoleucas* comenzó en el límite entre el Oligoceno Tardío y el Mioceno hace aproximadamente 30-20 Ma, cuando las conexiones terrestres entre Europa y Norte América todavía estaban presentes, y culminó tras la ruptura definitiva de estos hipotéticos puentes de tierra (Böhme, 2000; Imoto *et al.*, 2013).

Especialmente crítico fue el Mioceno (23-5,3 Ma), periodo en el que los eventos paleogeográficos fueron muy numerosos, muy rápidos y simultáneos a lo largo de toda la cuenca del Mediterráneo (Rögl, 1999). Divergencias ocurridas en el Mioceno Temprano (23-20 Ma) dieron lugar a la mayoría de los géneros actuales de la subfamilia Leuciscinae s.s., entre ellos *Squalius*, y coincide con la mayor actividad tectónica de la orogenia Alpina, que en resumen provocó el cierre del corredor de Eslovenia, una conexión marina entre el Tethys y el Paratethys, hace 20 millones de años; la extrusión y levantamiento de los Alpes del este; rotaciones de diversas microplacas; e inversión del terreno de la llanura Panónica, los Cárpatos y los Alpes Dináricos (Ratschbacher *et al.*, 1991; Frisch *et al.*, 1998; Paramonova *et al.*, 2004; Kuhlemann, 2007; Marton, 2007; Bartol *et al.*, 2008; Hinsbergen *et al.*, 2008). Estos eventos tectónicos, que en definitiva cerraron el corredor marino de Eslovenia y modificaron enormemente la paleogeografía de la región mediterránea, permitieron la diversificación de los principales grupos de leuciscinos en Europa (Perea *et al.*, 2010). La orogenia Alpina se extendió principalmente por todo el Cenozoico y dio lugar levantamiento de las principales cordilleras montañosas euroasiáticas, desde los Pirineos hasta el Himalaya (Muñoz *et al.*, 1986; López-Martínez, 1989; Puigdefábregas *et al.*, 1992; Paveliū y Belak, 2008), que actuaron como una importante barrera biogeográfica aislando la ictiofauna de las penínsulas mediterráneas de la del centro y norte de Europa, originando un alto grado de endemidad de la fauna dulceacícula en estas penínsulas (Šusnik *et al.*, 2001; Šlechtová *et al.*, 2004; Stefani *et al.*, 2004; Levy *et al.*, 2009; Perea *et al.*, 2010).

Un evento geológico ocurrido en el Mioceno y especialmente relevante para la ictiofauna continental de la región Mediterránea fue la apertura del mar Egeo, que se estima comenzó en el periodo Tortonense Tardío, hace aproximadamente 12 Ma, y terminó en el Serravaliense, entre 10-9 Ma (Rögl, 2001; Sprovieri *et al.*, 2003; Paramonova *et al.*, 2004; Çagatay *et al.*, 2006). Este

suceso supuso el surgimiento de linajes evolutivos vicariantes distribuidos a ambos lados del mar Egeo, entre ellos los pertenecientes al género *Squalius* (Doadrio y Carmona, 1998; Durand *et al.*, 2000; Kottelat y Economidis, 2006; Stomboudi *et al.*, 2006; Perea *et al.*, 2010). Sin embargo, estas relaciones vicariantes entre ambos lados del Egeo para el género *Squalius* son más complejas de forma que otros factores históricos deben estar implicados.

Tras el periodo de desecación del Mediterráneo al final del Mioceno (7-5,3 Ma), el Plioceno (5,3-2,6 Ma) se caracteriza por un aumento gradual de la humedad en la región mediterránea (Haywood *et al.*, 2016), hecho que pudo favorecer la diversificación de los leuciscinos en las penínsulas mediterráneas al disponer de nuevos ecosistemas de agua dulce potencialmente colonizables. Además, al final del Mioceno y durante el Plioceno se produjeron nuevos movimientos orogénicos que remodelaron la paleogeografía de las áreas adyacentes al plegamiento alpino (Popov *et al.*, 2006). Durante este periodo y hasta el Pleistoceno (2,6 Ma-10.000 a.C.), asociado a estas nuevas fases orogénicas alpinas, se fue configurando la red hidrológica actual de la región Mediterránea (Meulenkamp *et al.*, 2000; Popov *et al.*, 2004, 2006), y a su vez, siguiendo esta progresiva formación de las cuencas fluviales tuvieron lugar la mayoría de procesos de especiación que han originado la biodiversidad actual de fauna continental mediterránea al ir aislándose las nuevas cuencas formadas (ver por ejemplo Doadrio y Carmona, 1998; Zardoya y Doadrio, 1998, 1999; Englbrecht *et al.*, 2000; Perdices *et al.*, 2000, 2003; Sanjur *et al.*, 2003; Doadrio y Carmona, 2003, 2004, 2006; Kotlík *et al.*, 2004; Filipe *et al.*, 2009).

Más recientemente, las glaciaciones pleistocénicas, con ciclos alternando periodos glaciares y periodos interglaciares, han tenido una enorme influencia en la distribución de las especies de agua dulce de la región Mediterránea en los últimos dos millones de años de dos formas diferentes, en el centro y en el norte de Europa algunas de las poblaciones de estas especies se extinguieron localmente cuando las áreas que habitaban se cubrieron de hielo, mientras que las penínsulas mediterráneas actuaron como refugio (Banarescu, 1990, 1992; Hewitt, 1996, 1999, 2000, 2004; Reyjol *et al.*, 2007; Provan *et al.*, 2008). Por este motivo, la influencia de las glaciaciones pleistocénicas también se ha observado a escala poblacional, determinando enormemente la estructura filogeográfica de las poblaciones de peces de agua dulce, así como su diversidad genética asociada a la pérdida o ganancia de genotipos (Nesbø *et al.*, 1999; Kontula y Väinölä, 2001; Costedoat y Gilles, 2009; Hänfling *et al.*, 2002, 2009; Perdices *et al.*, 2003; Salzburger *et al.*, 2003; Hewitt, 2004; Kotlík *et al.*, 2004; Mäkinen y Merilä, 2008; DeFavari *et al.*, 2012; Vila *et al.*, 2016). Se ha propuesto una posterior recolonización de estas regiones del centro y norte de Europa a partir de refugios pleistocénicos de cuencas fluviales pontocáspicas (Mar Negro y Mar Caspio) en base a la amplia distribución actual de algunas

especies de ciprínidos como *Squalius cephalus* (L., 1758), *Chondrostoma nasus* (L., 1758) o *Alburnus alburnus* (L., 1758) (Durand *et al.*, 1999; Hewitt, 2000, 2004; Kotlík *et al.*, 2004; Kottelat y Freyhof, 2007a; Reyjol *et al.*, 2007; Gouskov y Vorburger, 2016). Sin embargo, estudios posteriores y nuestros propios datos sugieren patrones más complejos, como la presencia de refugios dentro de refugios y una diversidad críptica dentro de estas especies (Gómez y Lunt, 2007; Schmitt, 2007; Mäkinen y Merilä, 2008; Gante *et al.*, 2009; Seifertová *et al.*, 2012).

3. El género *Squalius* Bonaparte, 1837

3.1. *Biología y Ecología del género Squalius*

La familia Cyprinidae presenta un gran número de endemismos en la región Mediterránea (Kottelat y Freyhof, 2007a). Dentro de los ciprínidos, *Squalius* constituye uno de los géneros más diversificados y ampliamente distribuidos en la región Mediterránea y en Asia Occidental (Zardoya y Doadrio, 1999; Durand *et al.*, 2000; Sanjur *et al.*, 2003; Kottelat y Freyhof, 2007a; Perea *et al.*, 2010; Eschmeyer *et al.*, 2016). El género está formado por 52 especies formalmente descritas (Tabla 1), aunque se piensa que su diversidad específica está subestimada (Kottelat y Freyhof, 2007a; Eschmeyer *et al.*, 2016).

Existen numerosos trabajos específicos sobre la biología y ecología de algunas de las especies del género a una escala geográfica local (ver por ejemplo Ribeiro *et al.*, 2003; Maia *et al.*, 2006; Bostanci y Polat, 2009; Lorenzoni *et al.*, 2011; Gianetto *et al.*, 2013), algunos de estos estudios de índole ecológica tienen en cuenta interacciones bióticas con otras especies de peces de agua dulce (Sánchez-Hernández y Cobo, 2011; Logez *et al.*, 2013). Sin embargo, la información sobre la biología y ecología para algunas de las especies que han sido descritas recientemente es escasa, ya que además, muchas de estas nuevas especies descritas estuvieron previamente adscritas a otras especies de amplia distribución como *S. cephalus* o *S. lepidus* Heckel, 1843. Al igual que ocurre con la familia Cyprinidae considerada en conjunto, el género *Squalius* habita en una gran variedad de hábitats, desde ríos y arroyos de alta montaña en áreas de mayor o menor corriente, siendo algunas de las especies reófilas, hasta ecosistemas lénticos como lagos y embalses artificiales, preferentemente habitados por las especies de mayor tamaño (Kottelat y Freyhof, 2007a; Doadrio *et al.*, 2011; Caleta *et al.*, 2015). También están presentes en manantiales de sistemas kársticos, muy comunes en la región adriática (Caleta *et al.*, 2015). Las especies del género *Squalius* no presentan movimientos migratorios pre-reproductivos río arriba como ocurre con otros ciprínidos como barbos (*Barbus* y *Luciobarbus*) o bogas (*Chondrostoma* Agassiz, 1832 s.l.), pero realizan pequeños movimientos dentro del río en busca de pozas más profundas donde se refugian en invierno (Kottelat y

Freyhof, 2007a; Doadrio *et al.*, 2011). Su alimentación está basada principalmente en insectos, pero también pueden alimentarse de materia vegetal (Kottelat y Freyhof, 2007a). El hecho de que *Squalius* sea un género con un área de distribución tan amplia a lo largo de la región Mediterránea y Asia Occidental denota su potencial evolutivo para explotar multitud de nichos ecológicos, hecho que ha permitido la diversificación tan grande que ha tenido este grupo taxonómico.

El periodo reproductivo de la mayoría de las especies de *Squalius* ocurre después del primer o segundo año de vida y tiene lugar en primavera, entre febrero/marzo y junio (Kottelat y Freyhof, 2007a; Doadrio *et al.*, 2011; Caleta *et al.*, 2015). El género *Squalius* carece de un dimorfismo sexual conspicuo, aunque en la época reproductiva los machos desarrollan tubérculos nupciales (Miresan *et al.*, 2016). La fecundación es externa y suelen depositar la puesta sobre la grava; algunas especies presentan puestas fraccionadas, especialmente las especies que habitan en ecosistemas de tipología mediterránea (Kottelat y Freyhof, 2007a).

Respecto al estado de conservación del género *Squalius*, de las 52 especies descritas formalmente, 19 están incluidas en alguna categoría de amenaza de la UICN (www.iucnredlist.org): dos en la categoría de En Peligro Crítico (*Squalius cappadocicus* Özulu y Freyhof, 2011 y *Squalius palaciosi* Doadrio, 1980); 10 en la categoría de En Peligro y 7 en la de Vulnerable. Del resto de especies del género evaluadas por la comisión de expertos de la UICN, 5 especies figuran como Casi Amenazadas; 20 están incluidas en la categoría de Preocupación Menor y para dos de las especies se considera que existen Datos Insuficientes para realizar una evaluación de su riesgo de extinción. Algunas especies de *Squalius* también están incluidas en el Anexo III del Convenio de Berna y en el Anexo II de la Directiva Hábitats, como especies de interés comunitario que requieren la designación de áreas especiales para su conservación. Para muchas de las especies del género *Squalius* se ha observado un progresivo declive en su rango de distribución original a lo largo de toda la región Mediterránea, y esta reducción ha sido principalmente debida, entre otros factores, a la alteración de su hábitat y a la introducción de especies exóticas, una amenaza creciente a nivel global (Dudgeon *et al.*, 2006; Smith y Darwall, 2006; Doadrio *et al.*, 2011; Caleta *et al.*, 2015).

Tabla 1. Especies de *Squalius* descritas, localidades tipo y área de distribución.

Especie	Localidad tipo	Área de distribución
<i>Squalius adanensis</i> Turan, Kottelat y Doğan, 2013	Río Üçürge. Afluente de la cuenca del río Seyhan. Karaisali. Turquía	Cuenca del río Seyhan. Turquía
<i>Squalius agdamicus</i> Kamensky, 1901	Río Kura. Agdam. Azerbaijan	Cuenca del río Kura. Azerbaijan
<i>Squalius alburnoides</i> (Steindachner, 1866)	Afluente próximo a Mérida. España	Cuencas de los ríos Duero, Tajo, Sado, Guadiana, Odiel, Guadalquivir y Quarteira en la Península Ibérica
<i>Squalius albus</i> (Bonaparte, 1838)*	Lago Trasimeno. Italia	Lago Trasimeno. Italia
<i>Squalius anatolicus</i> (Bogutskaya, 1997)	Lago Beysehir. Turquía	Lagos Beysehir y Tuz. Turquía
<i>Squalius ahipsi</i> (Aleksandrov, 1927)**	Río Ahipsi. Afluente de la cuenca del río Kuban. Krepotsnaya. Rusia	Cuenca del río Kuban: afluentes entre los sistemas fluviales Agadum y Laba. Rusia
<i>Squalius aradensis</i> (Coelho, Bogutskaya, Rodrigues y Collares-Pereira, 1998)	Río Arade. Portugal	Ríos del sur de Portugal: Arade, Algibre y Borda
<i>Squalius aristotelis</i> Özuluğ y Freyhof, 2011	Río Behramkale. Península de Biga. Turquía	Cuenca del río Tuzla. Turquía
<i>Squalius berak</i> Heckel, 1843	Aleppo. Siria	Ríos de Siria, Turquía, Irán e Irak.
<i>Squalius cappadocicus</i> Özuluğ y Freyhof, 2011	Río Melendiz. Ilhara. Turquía	Lago Tuz. Turquía
<i>Squalius carinus</i> Özuluğ y Freyhof, 2011	Manatial Isikli. Turquía	Lago Isikli. Turquía
<i>Squalius carolitertii</i> (Doadrio, 1988)	Río Cega. Afluente de la cuenca del Duero. España	Ríos del cuadrante noroccidental de España y Portugal: Duero, Mondego, Limia, Miño, Otaivén y Léz. Algunas poblaciones en el río Tago
<i>Squalius castellanus</i> Doadrio, Perea y Alonso, 2007	Río Gallo. Afluente del río Tago. España	Cuenca del río Gallo, afluente de la cuenca alta del río Tago
<i>Squalius cephaloides</i> (Battalgil, 1942)	Armutlu. Turquía	Ríos del noroeste de Turquía
<i>Squalius cephalus</i> (Linnaeus, 1758)	Norte de Europa	Presenta una amplia distribución en Europa y Asia menor
<i>Squalius cii</i> (Richardson, 1857)	Río Gemlek. Bursa. Turquía	Isla de Lesbos en Grecia y ríos Bakir y Gediz en Turquía

Especie	Localidad tipo	Área de distribución
<i>Squalius fellowesii</i> (Günther, 1868)	Xanthos. Turquía	Cuencas de Grecia y Turquía alrededor del mar Egeo
<i>Squalius ghigii</i> (Gianferrari, 1927)	Isla de Rodas. Grecia	Isla de Rodas. Grecia
<i>Squalius illyricus</i> Heckel y Kner, 1857	Ríos Isonzo, Knin y Sign en la región de Dalmacia	Cuencas de los ríos Cetina y Krka en Croacia
<i>Squalius irideus</i> (Ladiges, 1960)	Marmaris Mugia-arasi. Turquía	Turquía: ríos entre las cuencas de Bergama y Dalaman
<i>Squalius janae</i> Bogutskaya y Zupancic, 2010	Río Dragonja. Península de Istria. Croacia	Cuenca del río Dragonja
<i>Squalius keadicus</i> (Stephanidis, 1971)	Río Evrotas. Peloponeso. Grecia	Cuenca del río Evrotas. Peloponeso. Grecia
<i>Squalius kosswigi</i> (Karaman, 1972)	Río Gümüldür. Izmir. Turquía	Cuenca del río Tahtali. Turquía
<i>Squalius kottelati</i> Turan, Yilmaz y Kaya, 2009	Embalse Tahtaköprü. Cuenca del río Orontes. Turquía	Turquía
<i>Squalius laietanus</i> Doadrio, Kottelat y de Sostoa, 2007	Río Matarraña. Nonaspe. España	Cuadrante nororiental de España, desde la cuenca del río Ebro al río Port Bou; en el río Ebro solamente se encuentra en la parte baja, principalmente en los ríos Segre y Matarraña. Cuencas de los ríos Tech y Agly en Francia.
<i>Squalius lepidus</i> Heckel, 1843	Río Tigris. Mosul. Irak	Cuencas de los ríos Ais Nehri, Quwayq, Tigris y Eufrates.
<i>Squalius lucumonis</i> (Bianco, 1983)	Río Ombrone. Italia	Cuencas de los ríos Serchio, Ombrone, Arno y Tiber en Italia
<i>Squalius malacitanus</i> Doadrio y Carmona, 2006	Río Guadalmina. Benahavis. España	Cuencas de los ríos Guadalmina, Guadaiza y Guadiaro en el sur de España
<i>Squalius microlepis</i> Heckel, 1843	Imotski. Bosnia-Herzegovina	Cuencas de los lagos Prološko y Riliče cerca de Imotski, cuenca de los ríos Matica y Tihaljina en Bosnia-Herzegovina
<i>Squalius moreoticus</i> (Stephanidis, 1971)	Lago Stymphalia. Peloponeso. Grecia	Lago Stymphalia y cuenca del río Vouraikos en Grecia

Especie	Localidad tipo	Área de distribución
<i>Squalius namak</i> Khaefi, Esmali, Sayyadzadeh, Geiger y Freyhof, 2016	Manantial Bolagh. Shazand. Irán	Lago Namak y cuenca del río Kavar en Irán
<i>Squalius orientalis</i> (Nordmann, 1840)	Abkhazia. Georgia	Cuencas de la vertiente del Mar Caspio
<i>Squalius orpheus</i> Kottelat y Economidis, 2006	Río Evros. Grecia	Cuencas entre los ríos Evros y Struma en Bulgaria, Grecia y Turquía
<i>Squalius palaciosi</i> (Doadrio, 1980)	Río Jándula. Cuenca del Guadalquivir. España	Tramo medio de la cuenca del río Guadalquivir
<i>Squalius pamvoticus</i> (Stephanidis, 1939)	Lago Pamvotis. Grecia	Lago Pamvotis y cuencas de los ríos Thyamis, Acheron, Louros y Arachthos en Grecia
<i>Squalius peloponensis</i> (Valenciennes, 1844)	Morea. Peloponeso. Grecia	Cuencas del Peloponeso griego excepto río Evrotas y lago Pamvotis
<i>Squalius platyceps</i> Zupancic, Maric, Naseka y Bogutskaya, 2010	Lago Skadar. Vranjina. Montenegro	Lago Skadar y cuenca del río Drin
<i>Squalius prespensis</i> (Fowler, 1977)	Lago Prespa y tributarios	Lago Prespa y tributarios
<i>Squalius pursakensis</i> (Hankó, 1925)	Kara-Chehir, Kötschke-Kissik y Eski-Chehir. Turquía	Cuenca del río Sakarya. Turquía
<i>Squalius pyrenaicus</i> (Günther, 1868)	Río Mondego y Sintra en Portugal	Mitad meridional de la Península Ibérica, desde el río Tago hacia el sur, con excepción de las cuencas próximas al Estrecho de Gibraltar. Presente en localidades puntuales en el río Ebro y en el tramo alto del río Júcar
<i>Squalius recurvirostris</i> Özüluğ y Freyhof, 2011	Arroyo en Ortaköy. Aksehir. Turquía	Cuencas de los lagos Eber, Aksehir y Ilgin. Turquía
<i>Squalius ruffoi</i> Bianco y Recchia, 1983	Río Svuto. Calabria. Italia	Cuencas del sur de Italia
<i>Squalius semae</i> Turan, Kottelat y Bayçeceli, 2017	Río Serçeme. Afluente del río Karasu. Erzurum. Turquía	Cuenca del río Karasu
<i>Squalius seyhanensis</i> Turan, Kottelat y Doğan, 2013	Río Satiz. Afluente del Seyhan. Turquía	Cuenca del río Seyhan
<i>Squalius spurius</i> Heckel, 1843	Aleppo. Siria	Turquía y Siria
<i>Squalius squalus</i> (Bonaparte, 1837)	Río Tiber. Italia	Se distribuye por la mayoría de las cuencas italianas y en la vertiente este del Adriático desde la cuenca del río Ofanto hasta el límite con las

Especie	Localidad tipo	Área de distribución
<i>Squalius svallize</i> Heckel y Kner, 1857	Lagos cerca de Vergoraz y Narenta	cuencas de los lagos Skadar y Ohrid. Cuencas de los ríos Neretva, Trebisnjica y Ljuta en Croacia y Bosnia-Herzegovina
<i>Squalius tenellus</i> Heckel, 1843	Livno. Bosnia-Herzegovina	Cuenca del río Cetina y en áreas kársticas de Livanjsko, lagos Busko y Madecko
<i>Squalius torgalensis</i> (Coelho, Bogutskaya, Rodrigues y Collares-Pereira, 1998)	Río Torgal. Odemira. Portugal	Cuenca del río Torgal en el sur de Portugal
<i>Squalius turcicus</i> De Filippii 1865	Río Arax. Erzurum. Turquía	Turquía e Irán
<i>Squalius valentinus</i> Doadrio y Carmona, 2006	Río Turia. Tuéjar. España	Vertiente mediterránea española, desde la cuenca del río Mijares al río Vinalopó
<i>Squalius vardarensis</i> Karaman, 1928	Yugoslavia	Desde la cuenca del río Sperchios a la cuenca del río Vardar en Grecia y Macedonia
<i>Squalius zrmanjae</i> Karaman, 1928	Tramo bajo del Río Zrmanja	Cuencas de los ríos Zrmanja y Krka en Croacia

* Algunos autores proponen la sinonimia de esta especie con *Squalius squalus* (Bianco, 2013). ** Algunos autores transfieren esta especie a *Petroleuciscus ahipsi* (Perea *et al.*, 2010).

*** Doadrio (1988) considera población del mondegó dentro de *S. carolitertii* y población de Sintra dentro de *S. pyrenaicus*.

3.2. Filogenia y Sistemática en el género *Squalius*

El género *Squalius* se caracteriza por los siguientes caracteres morfológicos: ausencia de una quilla medioventral delante de la región anal; línea lateral completa; boca generalmente terminal o subterminal; dos filas de dientes faríngeos (excepto *S. alburnoides*); origen de las aletas dorsales por detrás de la base de las aletas pélvicas; 8-11 radios ramificados en la aleta anal y de 7 a 9 en la aleta dorsal (Kottelat y Freyhof, 2007a). La sistemática del género *Squalius* es una cuestión pendiente por resolver a pesar de los grandes esfuerzos taxonómicos que se han realizado al respecto. El género *Squalius* fue previamente considerado en base a sus características morfológicas un subgénero dentro del género *Leuciscus*, que englobaba cuatro subgéneros diferentes: *Leuciscus* Cuvier, 1817; *Telestes* Bonaparte, 1837; *Squalius* Bonaparte, 1837; e *Idus* Heckel, 1858 (Berg, 1949; Banarescu, 1964). Sin embargo, algunos de los caracteres morfológicos y anatómicos que comparten entre sí y que los adscribían al género *Leuciscus* han resultado ser plesiomórficos para la familia Leuciscinae (Bogutskaya, 1990a,b; 1991, 1994). Por otro lado, la falta de un contexto filogenético para relacionar a estos cuatro subgéneros entre sí llevó a su posterior reconocimiento como distintos complejos de especies dentro del género *Leuciscus*, denominándose en el caso de *Squalius* “complejo *Leuciscus cephalus*” (Bianco, 1983). Finalmente, una vez conocido su marco filogenético, cada uno de estos subgéneros fue elevado al rango de género (Doadrio, 2001; Sanjur *et al.*, 2003).

Con 52 especies actualmente descritas (Tabla 1), el conocimiento de la diversidad del género *Squalius* casi se ha triplicado en los últimos veinte años de estudios sobre taxonomía, sistemática y filogenia del mismo (ver por ejemplo Doadrio, 1988; Doadrio *et al.*, 2003, 2004, 2006; Sanjur *et al.*, 2003; Stomboudi *et al.*, 2006; Kottelat y Freyhof, 2007; Turan *et al.*, 2007; Almada y Sousa-Santos, 2010; Perea *et al.*, 2010; Waap *et al.*, 2010; Özuluğ y Freyhof, 2011). Esto es debido en gran parte a los avances realizados en las últimas dos décadas en disciplinas como la Genética y la Sistemática Molecular, que han servido como complemento a la morfología clásica dentro de un pensamiento conceptual que tiende hacia una taxonomía integrativa (Dayrat, 2005; Schlick-Steiner *et al.*, 2010). El uso de herramientas procedentes de diversas disciplinas científicas en la sistemática de *Squalius* ha permitido un análisis más específico de la diversidad real del género y ha llevado a la identificación de linajes genéticos independientes y especies crípticas que estaban enmascaradas a consecuencia de la poca variabilidad morfológica externa que presenta el género *Squalius*. De esta forma las especies del género *Squalius* presentan escasa o nula variabilidad en algunos de los principales caracteres utilizados en la taxonomía de ciprínidos, como el número de radios ramificados en las aletas, el número de escamas, el número de branquiespinas o el número de vértebras (Doadrio, 1988; Doadrio y Carmona, 2006; Kottelat y Freyhof, 2007a; Özuluğ y Freyhof, 2011). Sin embargo,

algunos caracteres osteológicos permiten distinguir especies; entre estos caracteres cabe citar los huesos infraorbitarios o mandibulares (Fig. 3), muy diferentes entre especies cercanas como *S. carolitertii*, *S. castellanus* y *S. pyrenaicus* tratadas en el Capítulo II de esta tesis.

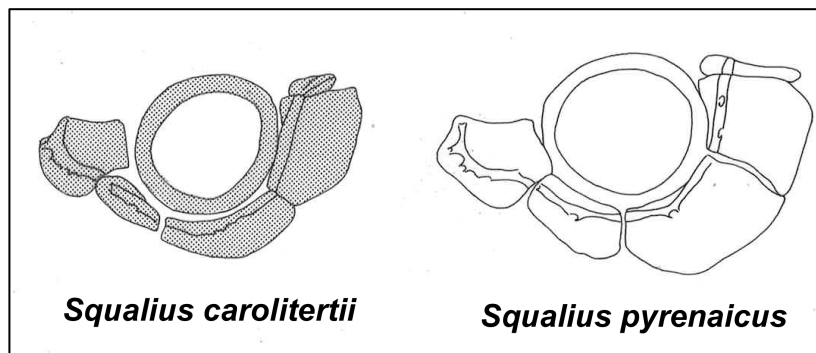


Fig. 3. Huesos circumorbitarios de *S. carolitertii* y *S. pyrenaicus*. En la figura se aprecia una mayor anchura de estos huesos en *S. pyrenaicus* respecto a *S. carolitertii*.

Respecto a su historia evolutiva, dentro del género *Squalius* se identificaron tres grandes linajes evolutivos (Sanjur *et al.*, 2003): un linaje “Mediterráneo” compuesto por especies de tamaño pequeño-mediano endémicas de las principales penínsulas mediterráneas, donde quedaron confinadas por fenómenos de especiación alopátrica como consecuencia del levantamiento de la Cordillera Alpina (Perea *et al.*, 2010); un segundo linaje “Euroasiático” formado por especies de mayor tamaño que el anterior y ampliamente distribuido por Europa central y parte de Asia, desde el río Ebro en la península ibérica hasta los montes Urales por el centro y norte de Europa y extendiéndose por el Cáucaso hacia la región del mar Caspio y Mesopotamia, sur de Irán y Siria; y, finalmente, un tercer linaje denominado “Paratethys” constituido principalmente por especies que habitan en cuencas hidrológicas alrededor del Mar Negro y en Anatolia. Este último linaje del “Paratethys” fue posteriormente asignado como un género distinto, *Petroleuciscus*, de acuerdo a su menor número de vértebras y de poros cefálicos y a diferencias osteológicas respecto a los representantes de los otros dos linajes de *Squalius* (Bogutskaya, 2002), siendo por tanto *Petroleuciscus* el grupo hermano del actual género *Squalius*. Un trabajo posterior ampliando el muestreo taxonómico de las especies del género corroboró la existencia de estos tres linajes evolutivos y determinó la adscripción a cada uno de estos tres linajes de especies de *Squalius* anteriormente no estudiadas desde un punto de vista molecular, así como una subestructuración geográfica en las especies del linaje “Mediterráneo” pero no en el “Euroasiático” (Perea *et al.*, 2010). El conocimiento de la diversidad de especies del linaje “Mediterráneo” y del linaje “Euroasiático” está muy desequilibrado, teniendo en cuenta la desproporción en el área de distribución de ambos linajes y el desconocimiento de la diversidad genética poblacional de algunas especies de amplia distribución como *S. cephalus*, *S. lepidus* o *S. orientalis*, pertenecientes al linaje “Euroasiático”.

La sistemática y la filogenia de las especies de *Squalius* pertenecientes al linaje “Mediterráneo” están razonablemente soportadas, aunque todavía quedan cuestiones pendientes por concluir. La mayor parte de estudios taxonómicos y evolutivos se han centrado en la península ibérica, donde se halla la mayor diversidad específica dentro de este linaje, ya que todas las especies ibéricas, excepto *S. laietanus*, pertenecen al linaje evolutivo “Mediterráneo” (9 especies; Kottelat y Freyhof, 2007a; Doadrio *et al.*, 2011). Por tanto, en la península ibérica son numerosos los trabajos científicos que han intentado dilucidar las relaciones filogenéticas de las especies de *Squalius*, tanto a nivel de género como a escala poblacional (Coelho *et al.*, 1995; Alves *et al.*, 1997; Brito *et al.*, 1997; Zardoya y Doadrio, 1998; Doadrio y Carmona, 2003, 2004; 2006; Sanjur *et al.*, 2003; Doadrio *et al.*, 2007a,b; Levy *et al.*, 2009; Almada y Sousa-Santos, 2010; Perea *et al.*, 2010; Waap *et al.*, 2010; Geiger *et al.*, 2014). Algunos de estos trabajos, sin considerar aquellos centrados en las especies de origen híbrido *S. alburnoides* y *S. palaciosi*, emparentadas mitocondrialmente con la especie *S. pyrenaicus* (Carmona *et al.*, 1997; Cunha *et al.*, 2004; Sousa-Santos *et al.*, 2007), se han fundamentado principalmente en analizar las relaciones filogenéticas del género *Squalius* desde un punto de vista mitocondrial, mientras que el uso de marcadores nucleares se ha restringido principalmente al estudio de alozimas e identificación de *loci* diagnósticos para las distintas poblaciones o especies y sólo unos pocos trabajos han intentado esclarecer las relaciones filogenéticas utilizando secuencias de ADN nuclear (Almada y Sousa-Santos, 2010; Perea *et al.*, 2010; Waap *et al.*, 2010).

El género *Squalius* en la península ibérica tiene un origen antiguo. Remontándonos cronológicamente en el tiempo, la especie fósil de ciprínido más antigua hallada en la península, *Rutilus antiquus* Cabrera y Gaudant, 1985, data del Oligoceno y fue descrita sobre la base de unos dientes faríngeos hallados en lo que fue la antigua cuenca terciaria del río Ebro. La especie actual más afín morfológicamente a *R. antiquus* es *S. alburnoides* (Cabrera y Gaudant, 1985), por lo que probablemente esta especie fósil también podría ser adscrita al género *Squalius*. Dos especies fósiles del Mioceno fueron identificadas como pertenecientes al género *Leuciscus*, del que *Squalius* fue previamente considerado un subgénero (Banareescu, 1964), *L. pachecoi* Royo 1922 del Turoliense inferior (Zona mastozoológica MN-11) y *L. antunesi* Gaudant, 1977 del Astaraciense (zona mastozoológica MN-7) (Cabrera y Gaudant, 1985; de la Peña, 1995). Por sus rasgos morfológicos, se ha considerado que la especie ibérica más próxima a *L. pachecoi* es *Achondrostoma arcasii* (Royo, 1921), previamente considerada una especie del género *Rutilus*, por lo que se transfirió *L. pachecoi* al género *Rutilus* (Gaudant, 1984; de la Peña, 1995). Sin embargo, por la morfología dentaria parece más probable que *L. pachecoi* pertenezca al género *Squalius* (Doadrio y Perea, 2005). De las especies actuales, el clado monofilético más divergente

de las especies Ibéricas de *Squalius* es el constituido por *S. aradensis* y *S. torgalensis*, especies que habitan en las pequeñas cuencas hidrográficas del sur de Portugal. Un segundo linaje monofilético estaría compuesto por dos especies, *S. carolitertii*, que se distribuye por el cuadrante noroccidental de la Península Ibérica y *S. castellanus*, que sólo se conoce en algunos afluentes de la cuenca alta del río Tago. Finalmente, un tercer linaje monofilético ibérico engloba tres especies, *S. pyrenaicus*, *S. malacitanus* y *S. valentinus*. En la presente tesis, los Capítulos I y II versan sobre las relaciones filogenéticas de las poblaciones de las especies de *Squalius* que habitan el sur de la península (*S. malacitanus* y *S. pyrenaicus*) y de las especies que presentan el área de distribución más amplia, *S. pyrenaicus* y *S. carolitertii*, con el fin de conocer su diversidad taxonómica y filogenética real.

En Italia solamente se cita un representante del linaje Mediterráneo, *S. lucumonis* (Bianco, 1983; Durand *et al.*, 2000; Sanjur *et al.*, 2003; Kottelat y Freyhof, 2007a; Perea *et al.*, 2010). En los Balcanes, la mayoría de las especies del linaje “Mediterráneo” se concentran en la región de Dalmacia, donde habitan las especies *Squalius microlepis*, *S. illyricus*, *S. svallize*, *S. tenellus* y *S. zrmanjae*, mientras que las otras dos especies balcánicas del linaje “Mediterráneo”, *S. keadicus* y *S. ghigii*, son endémicas del Peloponeso griego y de la isla de Rodas en Grecia respectivamente. Dentro de este linaje también se incluye la especie del este de Turquía *S. irideus*, (Kottelat y Freyhof, 2007a; Perea *et al.*, 2010; Caleta *et al.*, 2015). Las relaciones filogenéticas de las especies balcánicas mediterráneas han sido inferidas desde un punto de vista molecular (Durand *et al.*, 2000; Perea *et al.*, 2010; Geiger *et al.*, 2014), sin embargo, estos autores solamente incluyen uno o dos individuos por especie, por lo que no se conocen bien las relaciones filogenéticas y la estructura filogeográfica de las poblaciones de estas especies, como se muestra en el Capítulo III de esta tesis.

Respecto a las relaciones entre las especies de las tres penínsulas mediterráneas, las especies ibéricas constituyen un grupo monofilético independiente de las especies del resto de las penínsulas; la especie italiana *S. lucumonis* está emparentada con las especies de la región de Dalmacia, en la vertiente este del Mar Adriático, pero las especies griegas *S. keadicus* y *S. ghigii*, no forman parte del mismo clado, de hecho, la posición filogenética de estas especies respecto a las especies ibéricas o adriáticas del linaje “Mediterráneo” es incierta y variable dependiendo de los marcadores genéticos utilizados y de los análisis realizados (Durand *et al.*, 2000; Sanjur *et al.*, 2003; Perea *et al.*, 2010; Geiger *et al.*, 2014).

En el lado opuesto, tradicionalmente se consideraba que el linaje “Euroasiático” estaba compuesto por dos especies, *Squalius lepidus*, ubicada en la parte este de Anatolia, Mesopotamia y en el Levante Mediterráneo, y *S. cephalus* en el resto del área de distribución,

presentando ambas especies claras diferencias morfológicas (Berg, 1949; Bogutskaya, 1994, 1997). La especie *S. cephalus* fue dividida a su vez en *S. cephalus*, al oeste de su área de distribución, y *S. orientalis* al este (Berg, 1949), sin embargo, la frontera entre ambas taxa es difusa, fundamentalmente en la región oriental de la cuenca Mediterránea y en oriente próximo, donde se ha descubierto una variabilidad morfológica que no se corresponde con la división del linaje centroeuropeo en estas tres especies (Turan *et al.*, 2007; Özuluğ y Freyhof, 2011; Esmali *et al.*, 2016). Dentro del complejo de especies de *Squalius cephalus* se identificaron cuatro linajes genéticos en base a datos obtenidos a partir del gen mitocondrial citocromo *b* y de microsatélites, un linaje adriático, un linaje del Egeo, un linaje del oeste de Europa y un linaje del Este, que de acuerdo a estos autores descenderían de una rápida radiación pliocénica, hace aproximadamente tres millones de años, a partir de un ancestro común de origen mesopotámico (Durand *et al.*, 2000; Seifertová *et al.*, 2011). Además, datos moleculares también han demostrado la existencia de distintos grupos genéticos independientes dentro de estas tres especies (Doadrio y Carmona, 1998; Durand *et al.*, 2000; Sanjur *et al.*, 2003; Perea *et al.*, 2010; Geiger *et al.*, 2014), que junto con un análisis morfológico más exhaustivo de las especies euroasiáticas han derivado en la descripción de numerosas especies dentro de este linaje “Euroasiático” (Kottelat y Economidis, 2006; Stoumboudi *et al.*, 2006; Doadrio *et al.*, 2007a; Turan *et al.*, 2009, 2013, 2017; Bogutskaya y Zupančič, 2010; Zupančič *et al.*, 2010; Özuluğ y Freyhof, 2011; Kaefhi *et al.*, 2016). Aún así, la taxonomía y sistemática actual de los representantes de este género pertenecientes al linaje “Euroasiático” no es del todo bien conocida y no existe un marco filogenético actualizado para la mayoría de estas especies descritas en base a caracteres morfológicos, ya que los estudios analizando estas especies “euroasiáticas” desde el punto de vista molecular son escasos (Doadrio y Carmona, 1998; Durand *et al.*, 1999, 2000; Sanjur *et al.*, 2003; Doadrio y Carmona, 2003; Doadrio *et al.*, 2007a; Perea *et al.*, 2010). Sin embargo, por la extensión y complejidad de esta problemática, esta cuestión queda fuera del marco de la presente tesis.

Por otro lado, la adscripción al género *Squalius* de algunos taxa se ha cuestionado. Algunos estudios moleculares han propuesto la sinonimización del género *Ladigesocypris* con *Squalius* de acuerdo a la posición filogenética de dos especies del género (*L. ghigii* y *L. irideus*) inferida utilizando ADN mitocondrial y nuclear (Perea *et al.*, 2010). Aunque esta posición filogenética del género *Ladigesocypris* asociada a *Squalius* ya había sido inferida previamente en base a relaciones filogenéticas mitocondriales (Durand *et al.*, 2002; Sanjur *et al.*, 2003). Las especies *Ladigesocypris ghigii* y *L. irideus* fueron inicialmente descritas como *Leucaspius ghigii* Gianferrari 1927 y *Leucaspius irideus* Ladiges, 1960, posteriormente fueron transferidas al género *Ladigesocypris* y esta nomenclatura fue adoptada por diversos autores (Economidis, 1995; Bianco *et al.*, 1996; Bogutskaya, 1996, 1997; Kottelat, 1997; Kottelat y Freyhof, 2007a).

Algunos autores habían propuesto la inclusión del género *Ladigesocypris* en el género de ciprinido *Pseudophoxinus* en función de su morfología (Bogutskaya *et al.*, 2006), sin embargo, esta relación no tiene ningún soporte desde el punto de vista filogenético (Perea *et al.*, 2010). En esta tesis, el género *Ladigesocypris* se considera sinónimo de *Squalius*.

Por el contrario, se ha demostrado la adscripción molecular (mitocondrial y nuclear) de la especie *Squalius ahipsi*, previamente descrita como *Leuciscus ahipsi* Aleksandrov, 1927, al linaje genético del género *Petroleuciscus*, por lo que se ha defendido su transferencia a dicho género (Perea *et al.*, 2010), nomenclatura que se sigue en la presente tesis.

3.3. Patrones biogeográficos en el género *Squalius*

El género *Squalius* se distribuye a lo largo de la Región Mediterránea y Asia Occidental (Kottelat y Freyhof, 2007a; Doadrio *et al.*, 2011; Caleta *et al.*, 2015; Eschmeyer *et al.*, 2016). Hasta la fecha no existe ningún trabajo que abarque los patrones biogeográficos seguidos por el género *Squalius* al completo en toda su área de distribución. Algunos estudios exponen los patrones seguidos por este género en la región Mediterránea (Sanjur *et al.*, 2003) o dentro de un contexto biogeográfico general de la subfamilia Leuciscinae (Perea *et al.*, 2010), mientras que otros se centran en regiones concretas del área mediterránea con el fin de analizar patrones biogeográficos locales (Doadrio y Carmona, 1998; Sanjur *et al.*, 2003).

En base a un reloj molecular calibrado mediante el registro fósil se ha estimado que los géneros *Squalius* y *Petroleuciscus* divergieron en el Mioceno Inferior (hace aproximadamente 20-24 Ma) de su grupo hermano, constituido por un linaje que engloba a los actuales géneros *Chondrostoma s.l.*, *Phoxinellus*, *Telestes* y *Pseudophoxinus* (Perea *et al.*, 2010). Además, estos mismos autores estimaron que la divergencia entre los géneros *Squalius* y *Petroleuciscus* también tuvo lugar durante el Mioceno, hace aproximadamente 16-21 Ma. Los tiempos de divergencia de los distintos linajes evolutivos dentro de *Squalius* corresponden al Mioceno Superior, entre 7-11 Ma, de acuerdo a un reloj molecular mitocondrial calibrado con una tasa evolutiva de 0.76% por linaje por millón de años (Zardoya y Doadrio, 1998, 1999; Sanjur *et al.*, 2003). Desde entonces, el género *Squalius* se ha visto sometido a multitud de factores históricos (paleogeográficos y climáticos) que han determinado los patrones de distribución, asociados a procesos de especiación alopátrica, que observamos en la actualidad en este género. Tanto la divergencia de los géneros *Petroleuciscus* y *Squalius* como la separación de los dos linajes de *Squalius*, ocurridas en el Mioceno, coinciden con la mayor actividad tectónica de la orogenia Alpina y su influencia en la historia paleohidrológica de la región Mediterránea (Ratschbacher *et al.*, 1991; Frisch *et al.*, 1998; Paramonova *et al.*, 2004; Kuhlemann, 2007; Marton, 2007; Bartol *et*

al., 2008; Hinsbergen *et al.*, 2008). Las cadenas montañosas surgidas durante la orogenia Alpina supusieron barreras biogeográficas que aislaron los distintos linajes del género a medida que se fueron formando. Sin embargo, los procesos que dieron lugar a la diversificación de las especies, así como los factores históricos que las influenciaron se desconocen en la mayoría de los casos.

Al igual que ocurre con la filogenia y sistemática del género *Squalius*, la mayor parte de los estudios biogeográficos se han centrado en la Península Ibérica. El género *Squalius* en la península ibérica no constituye un grupo monofilético, sugiriendo dos eventos independientes de colonización por este género. La especie *S. laietanus*, endémica de la península ibérica y de la región suroccidental de Francia (principalmente en las cuencas de los ríos Agly y Tech), es filogenéticamente más próxima a las especies griegas y centroeuropeas del género (linaje “Euroasiático”), mientras que el resto de las especies, todas ellas endémicas de la Península Ibérica, se agrupan dentro del linaje “Mediterráneo” (Sanjur *et al.*, 2003; Doadrio y Carmona, 2003; Perea *et al.*, 2010; Denys *et al.*, 2013). Algunos de los modelos biogeográficos como la teoría dispersalista del “Lago Mare” se han refutado para el género *Squalius* en la península ibérica, ya que utilizando relojes moleculares calibrados se probó que el ancestro de las especies ibéricas de *Squalius* ya habitaba la península durante el Mioceno, antes de la fase lacustre del Mediterráneo, además el registro fósil apoya este resultado (Cabrera y Gaudant, 1985; de la Peña, 1995; Doadrio y Carmona, 2003; Perea *et al.*, 2010). Se ha estimado que las especies del sur de Portugal, *S. aradensis* y *S. torgalensis* divergieron del resto de especies ibéricas en el Messiniense, entre 7 y 5,3 Ma como consecuencia de la desecación y aislamiento de la mayor parte de las cuencas mediterráneas (Sanjur *et al.*, 2003; Mesquita *et al.*, 2007). Fenómenos de especiación asociados a la formación de cuencas sedimentarias endorreicas durante el Mioceno en la península ibérica y posterior transformación de estas cuencas al exorismo con la consecuente configuración de la red hidrográfica actual de la península han sido postulados para explicar la divergencia de especies como *S. carolitertii* y *S. pyrenaicus* (López-Martínez, 1989; Calvo *et al.*, 1993; Doadrio y Carmona, 2003, 2006; Sanjur *et al.*, 2003; De Vicente *et al.*, 2011; Pais *et al.*, 2012; Torne *et al.*, 2015), como también demostramos en los Capítulos I y II de esta tesis. Así, en los estos capítulos se investigan los patrones biogeográficos mostrados por varias especies del género *Squalius* pertenecientes al linaje “Mediterráneo” así como los factores históricos responsables de dichos patrones. Por el contrario, no existe hasta la fecha un modelo biogeográfico para explicar la presencia de la especie *S. laietanus* en la Península Ibérica, así como el hecho de que su grupo hermano lo formen las especies griegas del género.

En la región de los Balcanes considerada en su conjunto (países de la antigua Yugoslavia, Bulgaria y Grecia), existe un desequilibrio entre los estudios de biogeografía que se han realizado en Grecia para el género *Squalius* respecto al resto del territorio donde los

estudios biogeográficos para este género son escasos o incluso ausentes. El grado de endemismo de la fauna de los Balcanes es muy elevado (Oikonomou *et al.*, 2014). En Grecia se ha sugerido en base a un reloj mitocondrial estricto calibrado con la apertura del Estrecho de Gibraltar (5,3 Ma) que el origen de las especies del género *Squalius* y sus patrones de diversificación han tenido lugar a lo largo de los períodos Mioceno Medio y Plioceno Superior, entre 13 y 3 Ma (Doadrio y Carmona, 1998). En base a este reloj también se sugirió que la zona del Peloponeso griego, habitada por las especies *S. keadicus* (linaje “Mediterráneo”), *S. moreoticus* y *S. peloponensis*, pertenecientes estas dos últimas al linaje “Euroasiático”, se separó del resto de Grecia en el Mioceno Superior, hace al menos 5-10 Ma, aislando a *S. keadicus* del resto de especies griegas. Posteriores análisis incluyendo secuencias de ADN mitocondrial y nuclear, una calibración del reloj molecular con fósiles y un muestreo taxonómico más amplio, incluyendo no solamente las especies griegas del género, demostraron que los tiempos de divergencia para las especies griegas del género son más antiguos que los previamente propuestos (12-16 Ma; Perea *et al.*, 2010). La península balcánica está situada en el extremo sur de la placa euroasiática y es una zona muy compleja desde el punto de vista tectónico, principalmente el Peloponeso griego, que se ubica en el área próxima a la fosa oceánica Helénica en la zona de convergencia entre la placa euroasiática y la africana, donde todavía existen numerosas fallas activas (Kokkalas *et al.*, 2006; Papanikolaou *et al.*, 2007; Fountoulis *et al.*, 2008, 2014). Divergencias ocurridas en el Mioceno también se han obtenido para otras especies de ciprínidos endémicas del Peloponeso como *Tropidophoxinellus spartiaticus* (Schmidt-Ries, 1943) (Perea *et al.*, 2010). El levantamiento de la Cordillera del Pindo (Dermitzakis y Papanikolaou, 1981; Dermitzakis, 1990) también supuso una importante barrera biogeográfica implicada en la especiación alopátrica de las especies griegas “euroasiáticas” de *Squalius* durante el Mioceno Superior (Doadrio y Carmona, 1998).

Dos unidades geológicas significativas para la ictiofauna continental de los Balcanes son los sistemas de lagos Prespa y Ohrid, que se cuentan entre los lagos más antiguos de Europa con un origen datado en el Plioceno Inferior-Medio, hace 2-5 Ma cuando formaban parte de un lago de mayor tamaño denominado Dessateres (Stankovic, 1960; Meybeck, 1995; Wagner *et al.*, 2010). Ambos lagos se caracterizan por presentar un alto nivel de endemismo y singularidad en referencia a su ictiofauna respecto al resto de los Balcanes, hecho que ha suscitado que el lago Prespa sea reconocido además como una unidad biogeográfica independiente, sin embargo, la pertenencia o no del lago Ohrid a la misma unidad se ha cuestionado ya que aunque ambos sistemas lacustres tienen un origen común y están conectados entre sí a través de canales kársticos subterráneos la historia geológica de ambos es diferente y mientras el lago Ohrid ha mantenido un flujo constante de agua, el lago Prespa probablemente ha sufrido varios periodos de desecación a lo largo del tiempo (Amataj *et al.*, 2007; Zogaris *et al.*, 2009; Wagner *et*

al., 2010; Oikonomou *et al.*, 2014). A su vez, el lago Ohrid está conectado con el lago Skadar por el río Drim. El origen del lago Skadar es incierto, se considera que este lago en un principio fue una bahía marina que se rellenó de sedimentos fluviales de los ríos que actualmente lo alimentan y durante mucho tiempo más que un lago fue una laguna litoral (Boskovic *et al.*, 2004). El lago Prespa está habitado por la especie *S. prespensis* mientras que en el sistema de lagos Ohrid-Skadar se encuentra la especie *S. platyceps* (Doadrio y Carmona, 1998; Zupancic *et al.*, 2010), ambas pertenecientes al linaje “Euroasiático” de *Squalius* (Sanjur *et al.* 2003; Perea *et al.*, 2010) sin embargo, los patrones biogeográficos que han seguido los linajes de ambas especies en esta región biogeográfica no están claros.

3.4. Patrones filogeográficos e estructura intraespecífica en el género *Squalius*.

La filogeografía constituye una disciplina relativamente reciente. La palabra filogeografía fue acuñada en 1987 por Avise y colaboradores para definir los “principios y procesos que gobiernan la distribución geográfica de los linajes genealógicos, especialmente aquellos dentro y entre especies estrechamente emparentadas”; es decir, se encarga de analizar el “componente filogenético de la distribución espacial de los linajes de genes”, y por tanto tiene en consideración las variables tiempo y espacio (Avise *et al.*, 1987; Avise, 2000). La filogeografía pretende ser un puente entre la filogenia, la sistemática y la genética de poblaciones y ha sido comúnmente usada para evaluar la congruencia entre la distribución geográfica de las especies, o de sus poblaciones, y la estructura (o falta de estructura) genética de las mismas (Avise *et al.*, 1987; Avise, 1998, 2009; Moritz y Bermingham, 1998). La filogeografía también ha aportado un mejor entendimiento de procesos evolutivos recientes, como la influencia de las glaciaciones pleistocénicas sobre la formación de linajes genéticos y la distribución de los mismos o la identificación de refugios glaciares, así como la inferencia de procesos geológicos a partir del análisis de la estructura de las poblaciones (Waters *et al.*, 2001; Provan y Bennett, 2008).

Por sus características, que se describen brevemente en el epígrafe 5 de esta introducción general, el principal marcador genético utilizado en filogeografía de animales ha sido el ADN mitocondrial, el cual entre otros factores presenta una alta tasa evolutiva y una alta variabilidad en las secuencias nucleotídicas respecto al ADN nuclear, y por tanto es capaz de revelar en el caso de su existencia la presencia de una estructura genética intraespecífica y su asociación o no a una estructura geográfica (Avise, 2009). Sin embargo, los genes nucleares también son susceptibles de ser utilizados en aproximaciones filogeográficas siempre que cumplan una serie de requisitos, tales como que no presenten recombinación y que tengan la suficiente variabilidad para manifestar estructura poblacional, como por ejemplo ocurre con los intrones, regiones no codificantes, o los microsatélites (Hare, 2001; Avise, 2009).

Aunque en la península ibérica son numerosos los trabajos analizando la estructura filogeográfica de algunas de las especies diploides del género *Squalius*, especialmente de aquellas que presentan un área de distribución pequeña y restringida al sur de Portugal (Mesquita *et al.*, 2005, 2007; Henriques *et al.*, 2010), la estructura filogeográfica de algunas de estas especies en su áreas de distribución completa no se conoce bien, como es el caso de *S. carolitertii* y *S. pyrenaicus*, con una distribución transfronteriza entre los dos países que constituyen la península ibérica. Algunos estudios previos han demostrado cierta estructuración poblacional dentro de estas dos especies asociada a factores históricos de formación de las cuencas hidrológicas ibéricas durante los períodos Plioceno y Pleistoceno (Doadrio *et al.*, 2003, 2006; Sanjur *et al.*, 2003; Doadrio *et al.*, 2007; Almada y Sousa-Santos, 2010; Waap *et al.*, 2010). Los capítulos I y II de esta tesis exploran la estructura filogenética y filogeográfica de ambas especies en la península ibérica teniendo en cuenta su área de distribución completa.

El complejo de especies *Squalius cephalus* ha sido tradicionalmente el principal objetivo de los estudios filogeográficos ya que su estructura genética puede ser explicada en el contexto de las glaciaciones pleistocénicas. De esta forma, para este complejo se ha postulado la existencia de cuatro refugios pleistocénicos para esta especie (Adriático, mar Egeo, Danubio y cuencas del mar Caspio), a partir de los cuales se ha propuesto una colonización postglacial del centro y norte de Europa (Hänfling y Brandl, 1998; Durand *et al.*, 1999; Guinand *et al.*, 2001; Hewitt, 2004; Seifertová *et al.*, 2012; Gouskov y Vorburger, 2016). Sin embargo, las relaciones filogeográficas de este grupo son más complejas ya que su sistemática no está bien resuelta y algunos de los grupos considerados previamente *S. cephalus* ahora son reconocidos como especies distintas, ya que o bien se ha recuperado su denominación previa (por ejemplo *S. squalus* y *S. peloponensis*, que habían sido sinonimizadas con *S. cephalus* por Kottelat, (1997)), o bien se han descrito nuevas especies, como es el caso de *S. platyceps* o *S. laietanus* (Doadrio *et al.*, 2007a; Zupancic *et al.*, 2010).

4. Factores históricos climáticos que influyen en la historia evolutiva de los peces de agua dulce en la región Mediterránea: un resumen.

A lo largo de la historia de la Tierra diversos episodios climáticos han estado estrechamente ligados a los procesos evolutivos que han tenido lugar en la biota, en particular los peces de agua dulce. En la región Mediterránea dos son los principales eventos climáticos que han ejercido una fuerte influencia sobre las poblaciones de peces de agua dulce, el establecimiento del clima mediterráneo, con sus ciclos estacionales de lluvias/sequías y su impredecibilidad en cuanto a intensidad y duración de estos ciclos, y las glaciaciones pleistocénicas.

4.1. Origen del clima mediterráneo y consecuencias en la ictiofauna continental mediterránea

El clima mediterráneo se ha establecido en algunas regiones disyuntas del globo terráqueo (oeste de Norte América, sur de Australia, suroeste de Sudáfrica y Chile central, ocupando su mayor extensión en el territorio de los países situados en la cuenca del Mediterráneo) que presentan unas características fisionómicas, paisajísticas y ecológicas que no se encuentran en ningún otro lugar del mundo. Los sistemas acuáticos de las regiones de clima mediterráneo son ecológicamente únicos, con regímenes fluviales que se ajustan a los patrones estacionales e interanuales de precipitaciones (Gasith y Resh, 1999; Bonada y Resh, 2013). Aunque la ocurrencia de periodos de sequía y lluvias es predecible, no lo es su intensidad anual o su duración, afectando enormemente a las comunidades de agua dulce (Blondel y Aronson, 1999; Gasith y Resh, 1999; Bonada *et al.*, 2007; Di Castri y Mooney, 2012). La impredecibilidad estacional e interanual en la intensidad de los ciclos de lluvias/sequías de los ecosistemas acuáticos mediterráneos, junto con la progresiva explotación antropogénica de los mismos, ha provocado que estos ambientes se encuentren entre los más amenazados del mundo, especialmente importante a consecuencia de la creciente tendencia global hacia la desertificación (Kéfi *et al.*, 2007; Seager *et al.*, 2014). Sin embargo, la respuesta de la biota dulceacuícola ante esta tendencia progresiva a la aridez es distinta en función de su resistencia y adaptabilidad (concepto ecológico de resiliencia; Holling, 1973) a los cambios ambientales (Filipe *et al.*, 2013).

Para las comunidades de agua dulce de los sistemas acuáticos de regiones mediterráneas no solo es importante la media anual de precipitación sino también la distribución estacional de las precipitaciones, y sus ciclos biológicos se han adaptado a estas condiciones ambientales tan especiales. Durante los periodos de sequía, que suelen coincidir con la época estival, las condiciones ambientales no son las mismas que durante los periodos de lluvia y por tanto, la respuesta de los organismos vivos tampoco va a ser la misma en todas las épocas del año. En las épocas de sequía estival disminuye mucho el volumen de agua de los sistemas acuáticos a causa de una alta evapotranspiración, muchos de los ríos mediterráneos sufren un fuerte estiaje y su cauce queda reducido a pozas discontinuas, donde los organismos acuáticos quedan confinados hasta que las nuevas lluvias restauran la conectividad del cauce, lo cual puede ocurrir en un ciclo anual o en un ciclo interanual, no recobrándose la conectividad en una generación (Lake, 2003). Por este motivo, un elemento clave para la fauna acuática de sistemas mediterráneos es la presencia de refugios (Magoulick y Kobza, 2003). Esto provoca una reducción directa de hábitat, y por tanto de las áreas de puesta y reclutamiento, una

fragmentación de las poblaciones y que se exacerben las relaciones bióticas (competencia, depredación, alteración de las cadenas tróficas, etc) entre los componentes de las comunidades acuáticas, que a su vez condiciona enormemente la estructura demográfica de las poblaciones y, a largo plazo, la composición genética de las mismas (Frankham, 1996; Gasith y Resh, 1999; Lake, 2003; Ewers y Didham, 2006; Lowe y Allendorf, 2010). La estacionalidad mediterránea fomenta la fragmentación de las poblaciones y por tanto, se esperaría hallar una estructuración genética en las mismas; sin embargo, la restauración de la conectividad durante los periodos de lluvia puede diluir esta estructuración genética en mayor o menor medida al volver a conectar de nuevo las poblaciones (Hughes, 2007; Lowe y Allendorf, 2010). La fragmentación de poblaciones también se ha asociado con una pérdida de diversidad genética, una menor tasa de crecimiento y una disminución de la viabilidad poblacional (Frankham, 1996; Ewers y Didham, 2006). Si el estiaje es muy fuerte pueden incluso extinguirse poblaciones, como por ejemplo ocurre en zonas de tipología mediterránea de regiones semidesérticas australianas donde la sequía puede durar años, con la consecuente pérdida de genotipos y afectación a la diversidad genética (Spielman *et al.*, 2004; Faulks *et al.*, 2010). Por tanto, la estructura y diversidad genética de las poblaciones de peces de agua dulce de la región mediterránea va a estar enormemente influenciada por los ciclos de sequía/lluvias y su dinámica fluctuante a lo largo del tiempo. Existen multitud de ejemplos analizando a influencia del clima Mediterráneo sobre la estructura y diversidad genética de las poblaciones de peces de agua dulce (ver entre ellos Faulks *et al.*, 2010, 2011).

El clima Mediterráneo lleva millones de años instaurado en las regiones de tipología mediterránea, pero no es hasta el final del Plioceno, hace aproximadamente 3,2 Ma, cuando estos ciclos de sequías en periodos cálidos y lluvias en periodos templados-fríos se establecieron definitivamente, con una modificación sustancial de los ecosistemas mediterráneos tanto terrestres como acuáticos (Suc, 1984; Fauquette *et al.*, 1999; Jiménez-Moreno *et al.*, 2010; Haywood *et al.*, 2016). Estos cambios climáticos que ocurrieron en las áreas actuales mediterráneas están relacionados con cambios de temperatura en latitudes más septentrionales, pero el comienzo de la evolución del clima mediterráneo en estas áreas estuvo más relacionado con modificaciones en la humedad ambiental, especialmente en los ritmos de precipitaciones, que con cambios de temperatura propiamente dichos (Suc y Zagwijn, 1983; Suc, 1984).

En el Mioceno Medio, hace aproximadamente 18-16 Ma, se produjo un óptimo climático que se tradujo en un aumento de la temperatura y la humedad respecto a periodos anteriores más fríos y secos y es consistente con un aumento en la región Mediterránea de las floras termófilas y los animales ectotermos (Zachos *et al.*, 2001; Böhme, 2003). Tras este periodo se han registrado varios episodios de estacionalidad en las precipitaciones en un periodo más reciente dentro del Mioceno (16-14 Ma), pero no así en las temperaturas, que se mantuvieron

prácticamente sin cambios hasta el final del Mioceno Medio (14-13 Ma), cuando se produjo un descenso muy acusado de las temperaturas medias en algunas regiones del continente europeo, marcando el comienzo de una zonificación climática en la región Mediterránea respecto a áreas más septentrionales (Böhme, 2003). A partir de entonces, se produjo una tendencia creciente hacia condiciones más áridas y hacia una estacionalidad con períodos secos y lluviosos hasta que en el final del Plioceno surgió una estacionalidad en las temperaturas que se superimpuso a la preexistente estacionalidad en las precipitaciones, emergiendo el clima Mediterráneo tal como lo conocemos hoy día (Van Dam y Weltje, 1999; Jiménez-Moreno *et al.*, 2010).

En el Capítulo IV de esta tesis se evalúa la respuesta evolutiva de la especie *Squalius valentinus* ante la aparición y establecimiento del clima mediterráneo en el Plioceno a través del estudio de su estructura y diversidad genética (Fauquette *et al.*, 1999), a través de la evaluación de su influencia en la estructura y diversidad genética de la especie. La especie *S. valentinus* es endémica del Levante español y por tanto constituye un modelo de estudio adecuado para responder a este tipo de cuestiones evolutivas en una región donde la estacionalidad mediterránea puede ser extrema (Lázaro *et al.*, 2001).

4.2. Glaciaciones pleistocénicas y consecuencias en la ictiofauna continental mediterránea

Uno de los procesos climáticos más significativos que ha influido en la biota de la región Mediterránea en los dos últimos millones de años fueron las glaciaciones pleistocénicas. Este es un periodo caracterizado por la presencia de fuertes fluctuaciones climáticas en las que se alternan periodos glaciares, (fríos y secos) con periodos interglaciares, más cálidos y húmedos (Hays *et al.*, 1976; Paillard, 1998). La duración de estos periodos fue variable. Durante el Último Máximo Glacial la temperatura oscilaba alrededor de 21°C por debajo de la actual, por el contrario, durante algunos de los periodos interglaciares la temperatura fue más alta que en la actualidad (Cuffey *et al.*, 1995; Siroco *et al.*, 2005). El “secuestro” del hielo en glaciares se tradujo en un descenso en el nivel global del mar. Esto a su vez provocó una retracción en la línea de costa y por tanto un aumento de la superficie global de tierras emergidas, que pudo dar lugar a conexiones entre cuencas hidrológicas por sus desembocaduras (Alley *et al.*, 2005; Miller *et al.*, 2005; Sathiamurthy y Voris, 2006).

Estos eventos climáticos pleistocénicos dejaron una señal en los genomas de los organismos vivos, pero el efecto de las glaciaciones pleistocénicas sobre la biota dulceacuícola no fue homogéneo a lo largo de toda la región Mediterránea, más teniendo en cuenta que las especies de agua dulce poseen constricciones asociadas a su confinamiento a las cuencas

hidrográficas y a la historia geológica de la región que habitan (Myers, 1938; Hewitt, 2004). Durante los periodos glaciares gran parte del centro y norte de Europa se cubrió de hielo, de forma que algunas poblaciones se extinguieron localmente, dando lugar a pérdida de genotipos, y otras actuaron como refugios glaciares, funcionando como reservorios de variabilidad y variación genética, siendo para la fauna de peces continental las penínsulas mediterráneas y las cuencas pontocáspicas los refugios más característicos (Hewitt, 2000; 2004; Hänfling *et al.*, 2002, 2009; Kotlík *et al.*, 2004; Schmitt, 2007). Como se ha mencionado previamente en el epígrafe 3.2 de esta tesis, en realidad los patrones de refugios fueron más complejos de lo tradicionalmente considerado (refugios dentro de refugios; Gómez y Lunt, 2007; Gante *et al.*, 2009; Schmitt, 2007).

En las penínsulas mediterráneas más que extinción local de poblaciones asociada a los ciclos de glaciación/deglaciación se produjo una alteración en la dinámica fluvial, ya que las placas de hielo no llegaron a cubrir totalmente estas áreas meridionales pero sí que se vieron afectadas por el deshielo de los glaciares presentes en los sistemas montañosos. De la misma forma, los glaciares dieron lugar a cambios eustáticos en el nivel del mar, especialmente durante el Último Máximo Glacial, cuando el nivel del mar alcanzó su valor más bajo, 120-135 m por debajo del nivel actual (Alley *et al.*, 2005; Miller *et al.*, 2005; Hughes *et al.* 2006, 2007; Hughes y Woodward, 2008; Domínguez-Villar *et al.*, 2013; Adamson *et al.*, 2014). Por tanto, en las regiones del sur de Europa las poblaciones de peces continentales pudieron sufrir fuertes cambios demográficos, como cuellos de botella, reducción del tamaño efectivo o expansiones poblacionales dependiendo del régimen hídrico causado por el deshielo. Además, las potenciales conexiones entre cuencas hidrográficas, tanto por el descenso del nivel del mar como por fenómenos provocados por las grandes avenidas en épocas de deshielo pudo significar la dispersión y mezcla entre poblaciones de peces fluviales (Swartz *et al.*, 2007; Ruzzante *et al.*, 2008; Elmer *et al.*, 2009; Unmack *et al.*, 2011). Estas variaciones demográficas acontecidas durante las glaciaciones pleistocénicas en las penínsulas mediterráneas deben dejar una huella genética en las poblaciones que nosotros podremos descifrar para reconstruir la historia genética de las mismas en los últimos casi dos millones de años de la historia de la Tierra (Excoffier *et al.*, 2009).

En los capítulos I y IV de esta tesis se evalúa a través de reconstrucciones batimétricas el efecto de las glaciaciones pleistocénicas sobre la conectividad de las cuencas hidrológicas como consecuencia de cambios eustáticos en el nivel del mar durante el Último Máximo Glacial. Por tanto, indirectamente se analiza la conectividad de las poblaciones de peces de agua dulce a través de estas conexiones potenciales por los tramos bajos de los ríos y su impacto en la estructura y diversidad genética de estas poblaciones.

5. Marcadores moleculares y métodos de análisis utilizados para estudiar la historia evolutiva de los ciprínidos.

5.1. Marcadores moleculares: ventajas e inconvenientes

El descubrimiento de la estructura de la doble hélice de ADN en la segunda mitad del Siglo XX por Watson y Crick (1953), gracias en parte a los trabajos de difracción de rayos X de Rosalind Franklin (1950) supuso el comienzo de una revolución en la Biología Molecular y la Genética que ha ido creciendo a lo largo del tiempo hasta llegar a los avances que nos brindan hoy los campos de la Genómica y Transcriptómica en el contexto de la secuenciación de nueva generación (Schuster, 2008). El uso de marcadores moleculares en la historia evolutiva de los seres vivos comenzó en los años 70, especialmente desde el desarrollo de la técnica de Reacción en Cadena de la Polimerasa (PCR) y la secuenciación de Sanger (Sanger y Coulson, 1975; Sanger *et al.*, 1977; Kocher y White, 1989; Kocher *et al.*, 1989). Estas técnicas supusieron un gran avance metodológico y con el impulso y desarrollo de las mismas se ha producido un enorme avance en los estudios evolutivos a través de marcadores moleculares.

Para desentrañar la historia evolutiva de un taxón de forma precisa es conveniente el estudio de marcadores mitocondriales y nucleares, y por ello en la actualidad las aproximaciones multilocus para obtener una hipótesis filogenética son cada vez más frecuentes por su fiabilidad y robustez (ver por ejemplo Pereira *et al.*, 2002; Egger *et al.*, 2007; Dupuis *et al.*, 2012). Los análisis basados en un sólo tipo de marcador pueden dar lugar a interpretaciones erróneas de los patrones filogenéticos debido a procesos como coalescencias profundas (separación incompleta de linajes) o hibridación, como se espone en el epígrafe 5.4 de esta introducción general.

Uno de los marcadores moleculares más extensamente utilizado en los estudios evolutivos es el ADN mitocondrial, ya sea a través del uso de mitogenomas completos o de genes independientes. El genoma mitocondrial es una molécula circular de doble banda de típicamente 13-20 Kb constituida por 13 genes codificantes, relacionados con la cadena respiratoria, y 22 genes tRNA y dos genes rRNA relacionados con la traducción de proteínas, con algunas excepciones a esta organización general (Taanam, 1999; Lynch, 2007). El genoma mitocondrial presenta una serie de ventajas que lo hacen apto para ser utilizado en estudios evolutivos de los seres vivos (Avice *et al.*, 1987, 2009; Kocher *et al.*, 1989; Palumbi *et al.*, 1991; Moore, 1995; Ballard y Rand, 2005). Entre estas ventajas se citan las siguientes: i) su aislamiento en el laboratorio está plenamente estandarizado y existen multitud de

oligonucleótidos universales para la amplificación de sus genes por PCR para muchos grupos taxonómicos; ii) es haploide y carece de estructuras presentes en el ADN nuclear (estructuras repetitivas, transposones e intrones) que dificultan su análisis; iii) es de herencia materna por lo que no presenta recombinación; iv) tiene una tasa evolutiva relativamente rápida respecto al ADN nuclear al presentar múltiples copias por mitocondria (por lo que aumenta la probabilidad de mutación). Todo ello sumado al menor tamaño efectivo del ADN mitocondrial respecto al nuclear, que favorece la fijación de mutaciones con mayor probabilidad que en el ADN nuclear, permite inferir relaciones filogenéticas entre especies estrechamente emparentadas al reducir la discordancia debida a una separación incompleta de linajes. Además, la mayoría de los individuos son homoplásmicos para el ADN mitocondrial, es decir, todas las copias de sus células son idénticas; la heteroplasmia se da en baja frecuencias en animales, y cuando existe, las moléculas de ADN mitocondrial sólo suelen mostrar unas pocas mutaciones nucleotídicas (Avise, 2009). Otra ventaja del uso de genes mitocondriales es que evita el problema de paralogía, ya que además es de copia única, y elimina la ambigüedad procedente de los alelos heterocigotos (Li *et al.*, 2007; Chen *et al.*, 2008; Saitoh y Chen, 2008).

En los peces de agua dulce los dos genes mitocondriales más utilizados en estudios evolutivos son el gen que codifica para la enzima *citocromo b* (abreviado *MT-CYB* en la presente tesis) y el gen que codifica para la enzima *citocromo c oxidasa subunidad I* (abreviado como *COI* en la presente tesis). Aunque otros genes mitocondriales como los que codifican para diversas enzimas *nitrógeno deshidrogenasas* (*ND2*, *ND4* y *ND5*) o los genes que codifican para las enzimas *ATPasa 6* y *ATPasa 8*, así como la región control (no codificante) también han sido ampliamente utilizados en estudios filogenéticos y filogeográficos de peces de agua dulce (ver por ejemplo Zardoya y Meyer, 1996; Zardoya y Doadrio, 1998; Machordom y Doadrio, 2001a,b; Xiao *et al.*, 2001; Doadrio *et al.*, 2003; Sanjur *et al.*, 2003; Mäkinen y Merilä, 2008; Almada y Sousa-Santos, 2010; Perdices *et al.*, 2016).

El gen *MT-CYB* se encuentra en la mitocondria de todos los organismos eucariotas y en muchos procariotas, indicando un origen muy antiguo. Debido a que el *MT-CYB* presenta regiones altamente conservadas y regiones con una variabilidad considerable es un gen adecuado para responder cuestiones evolutivas, tanto entre grupos taxonómicos estrechamente emparentados como entre aquellos que son algo más distantes (Lydeard y Roe, 1997). El gen *MT-CYB* ha sido incluido como uno de los marcadores principales en las filogenias y filogeografías realizadas a lo largo de esta tesis debido a que ha sido un marcador extremadamente robusto en casi todos los grupos taxonómicos de peces de agua dulce (Kocher y Stepien, 1997).

El gen *COI* fue designado como un candidato idóneo para establecer un código de barras genético universal que permitiese identificar de forma diagnóstica, rápida y sencilla a las especies, como literalmente exponen los autores que seleccionaron un fragmento del gen *COI* de 655 pares de bases para tal fin (“...serve as the core of a global bioidentification system for animals”; Hebert *et al.*, 2003). De esta forma, el código de barras genético universal se planteó como apoyo a la taxonomía y sistemática ante la gran diversidad de organismos vivos y la creciente disminución de taxónomos para describir toda la biodiversidad existente (Hebert *et al.*, 2003; Hebert y Gregory, 2005; Tautz *et al.*, 2003). Este código de barras genético también ayudaría en casos de dificultad de identificación morfológica (especies crípticas, etc), e incluso tendría una finalidad más práctica para identificación y discriminación de especies en actividades socioeconómicas como la alimentación (por ejemplo stocks reproductivos en pequerías), de forma que actualmente son muy numerosos los trabajos que se han publicado al respecto en muchos grupos de vertebrados e invertebrados y en diversas regiones del mundo (ver entre otros Vences *et al.*, 2005; Ball y Armstrong, 2006; Robba *et al.*, 2006; Ward *et al.*, 2007, 2008; Geiger *et al.*, 2014). Sin embargo, a su vez, la adecuada aplicación de un código de barras depende de una correcta identificación de los organismos implicados en el mismo.

En el caso de los peces, algunos autores han defendido la potente capacidad de discriminación del código de barras basado en el gen *COI* y su utilidad para poner de manifiesto conflictos taxonómicos, incluso en áreas megadiversas como los trópicos (ver por ejemplo Ward *et al.*, 2007; Lara *et al.*, 2009; Pereira *et al.*, 2011, 2013). Por el contrario, otros han demostrado algunas limitaciones de este gen en condiciones especiales, como ocurre en los peces de agua dulce en la región Mediterránea a causa de su heterogeneidad espacial, o cuando las especies tienen un origen reciente (Tautz *et al.*, 2003; Geiger *et al.*, 2014). Así mismo, un código barras genético basado en un gen mitocondrial no puede identificar fenómenos de hibridación entre especies, por lo que puede llevar a resultados erróneos cuando estos procesos evolutivos están implicados.

Por su parte, el genoma nuclear de los animales contiene al menos 100 Megabases ordenadas en más de 13.000 genes, los cuales se organizan a su vez en cromosomas lineales, aunque estos valores varían entre especies (Lynch, 2007). Tal cantidad de material genético confiere al ADN nuclear la ventaja de poder muestrear loci independientes, incrementando el potencial para reconstruir hipótesis filogenéticas robustas (Page y Holmes, 1998). En los genes nucleares pueden distinguirse dos regiones distintas, los exones, que constituyen los fragmentos de los genes que contienen la información codificante para proteínas, y los intrones, que se intercalan entre los exones y forman parte de la transcripción primaria de ARN pero que son eliminados del transcrito maduro previamente a su traducción a proteínas (Lynch, 2007). A

efectos de su uso en análisis filogenéticos y filogeográficos, los intrones presentan una variabilidad mayor y una tasa de mutación más alta que las regiones codificantes, pero el alineamiento de los intrones a veces puede ser difícil, especialmente en análisis comparativos de grupos taxonómicos muy alejados (Chen *et al.*, 2008). Por este motivo es necesario llegar a un compromiso entre facilidad de trabajar y suficiente señal filogenética de las diferentes regiones de ADN para responder la pregunta evolutiva planteada.

Por su tasa evolutiva más lenta, especialmente en aquellas regiones que contienen la información codificante a proteínas, el ADN nuclear es adecuado para resolver relaciones filogenéticas en niveles taxonómicos altos (familias, órdenes, etc) (Vawter y Brown, 1986; Avise, 2009). Sin embargo, los exones suelen ser menos adecuados para relaciones más recientes, o para filogeografías a nivel poblacional. Ello es debido a que las relaciones pueden verse oscurecidas por procesos de separación incompleta de linajes porque no haya pasado el suficiente tiempo evolutivo para que las especies o las poblaciones hayan divergido respecto al marcador nuclear analizado (Maddison, 1997; Maddison y Knowles, 2006). De hecho, cuando se utilizan algunos determinados marcadores nucleares pueden no presentar la suficiente señal filogenética para resolver las relaciones a nivel de especie en determinados grupos taxonómicos (Springer *et al.*, 2001; Ballard y Rand, 2005; Wiens *et al.*, 2008; Fisher-Reid y Wiens, 2011), aunque esto también puede ocurrir con marcadores mitocondriales como consecuencia de saturación en los nodos profundos (ver por ejemplo Talavera y Vila, 2011; Mendes *et al.*, 2016). No obstante, en la gran variedad de genes nucleares existentes el grado de variabilidad en la secuencia de nucleótidos y la tasa de sustitución es muy variable debido a diferentes presiones de selección de acuerdo a la diversidad de funciones que muestran (Hillis, 1987). Además, los análisis han mejorado enormemente para lidiar con estos problemas introduciendo algoritmos de coalescencia, como se muestra en el apartado 5.3 de la presente tesis.

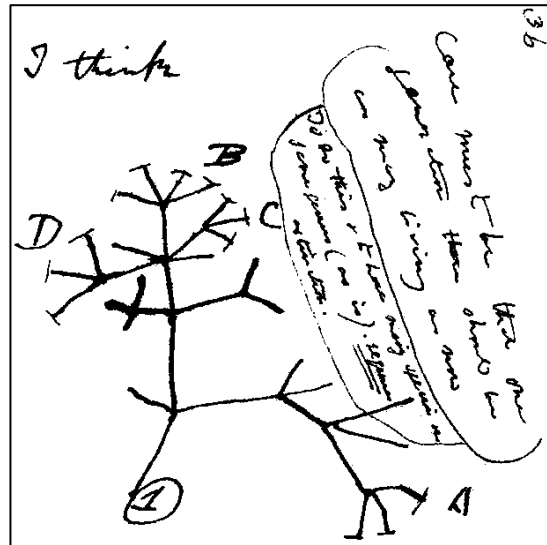
Sin embargo, el ADN nuclear tiene la ventaja de presentar menor homoplasia que el ADN mitocondrial (minimiza la saturación) y de permitir reconocer señales de introgresión e hibridación como consecuencia de recombinación de alelos, que no son posibles de observar si solamente se emplean marcadores mitocondriales en las filogenias (Ballard y Rand, 2005; Fisher-Reid y Wiens, 2011). En el genoma nuclear la recombinación entre *loci* homólogos durante la meiosis es frecuente, de esta forma la historia genealógica de un gen individual puede no coincidir con la historia evolutiva del grupo (Lynch, 2007). Esto es debido a que existan fenómenos de hibridación entre taxa o a la retención de polimorfismos ancestrales al azar dentro de un mismo taxón, especialmente en poblaciones con tamaños efectivos grandes (de nuevo separación de linajes incompleta; Maddison, 1997; Madison y Knowles, 2006). En el Capítulo V de la presente tesis se muestra un claro ejemplo de captura mitocondrial entre dos especies del

género *Squalius* en el Peloponeso griego que no hubiera sido detectada si no se hubiesen analizado los correspondientes genes nucleares de las especies implicadas en el fenómeno de introgresión descrito.

Por otro lado, un problema añadido puede ser el trabajar con genes nucleares que presentan varias copias a lo largo del genoma, ya que esto dificulta el aislamiento y la identificación de las copias ortólogas (aquellas procedentes de parentesco filogenético) imprescindibles para la obtención de una filogenia fiable (Avice, 2009; Poczai y Hyvönen, 2010). Mayor complejidad para identificar las copias ortólogas todavía viene dada en organismos poliploides (Comber y Smith, 2004; Otto, 2007). Los seis genes nucleares que se utilizan a lo largo de esta tesis, cinco codificantes (RAG1, S7, Rodopsina, EFA1 α , EGR2b y β -actina) y el primer intron del gen ribosomal S7 (no codificante), son genes de copia única en el genoma nuclear y han sido ampliamente utilizados en estudios de diferentes grupos taxonómicos de peces de agua dulce para resolver cuestiones filogenéticas, biogeográficas y filogeográficas de forma robusta (ver por ejemplo Hrbek *et al.*, 2007; Chen *et al.*, 2008; Říčan *et al.*, 2008; Bufalino y Mayden, 2010a,b; Marková *et al.*, 2010; Domínguez-Domínguez *et al.*, 2011; Unmack *et al.*, 2011; Yang *et al.*, 2015; Perdices *et al.*, 2016), entre ellos en el género *Squalius* (Almada y Sousa-Santos, 2010; Perea *et al.*, 2010; Waap *et al.*, 2010).

5.2. Filogenia Molecular

La variación fenotípica de los pinzones de las islas Galápagos inspiró a Darwin para formular su teoría de la evolución y su idea de árbol de la vida, tal como expone en su libro *El origen de las especies* (1859): “The affinities of all beings of the same class have sometimes been represented by a great tree...”. De esta forma, la evolución puede ser definida como descendencia con modificación desde un ancestro común y representada como un árbol filogenético ya que es la mejor manera de entender y describir las relaciones filogenéticas de los distintos grupos taxonómicos, y como estos se han ido diversificando a partir de un antepasado común (Fig. 3). Al final lo que se obtiene es una hipótesis de una historia evolutiva para un grupo taxonómico determinado, es decir, la topología del árbol refleja las relaciones evolutivas entre los componentes del árbol y como evolucionaron dichas relaciones.



La inferencia filogenética es un procedimiento de estimación estadística, ya que la información sobre la que se trabaja es incompleta, pues normalmente solo se tienen datos de los grupos taxonómicos actuales y aunque para algunos grupos se conocen fósiles su posición en el árbol filogenético a veces es incierta.

Los métodos de reconstrucción de árboles pueden subdividirse en aquellos que están basados en agrupamiento (métodos de distancias) y en aquellos que están basados en estados de caracteres y que utilizan un criterio de optimización para reconstruir el mejor árbol de entre todos los posibles (máxima parsimonia, máxima verosimilitud e inferencia bayesiana). Los métodos de agrupamiento definen una serie de pasos que conducen a la determinación de un único árbol. Para reconstruir el árbol filogenético, los métodos basados en distancias primero construyen una matriz de distancias pareadas entre todas las secuencias del alineamiento (nucleótidos o aminoácidos) y después estiman el árbol a partir de dicha matriz buscando la topología con la menor longitud, es decir, aquella en la que la distancia es mínima (Cavalli-Sforza y Edwards, 1967; Saitou y Nei, 1987; Swofford *et al.*, 1996). En el caso de secuencias de ADN la distancia genética entre dos pares de secuencias se calcula en base al número total de sustituciones de bases nitrogenadas (Saitou y Nei, 1987). Los métodos de distancia tienen algunas críticas, ya que solo muestran el grado de similitud entre las secuencias analizadas y no relaciones evolutivas, por tanto no pueden considerarse un verdadero método filogenético, además, como solamente producen un árbol estos métodos no proporcionan un apoyo estadístico a las relaciones inferidas (Swofford *et al.*, 1996; Yang y Rannala, 2012). Dos de los métodos de agrupamiento más comúnmente utilizados son UPGMA y unión del vecino (*Neighbor Joining*). UPGMA es la abreviatura de *Unweighted Pair Group Method with Arithmetic means*, en este método el agrupamiento de secuencias se realiza buscando el valor más bajo en la matriz

de distancias para formar un cluster inicial, posteriormente se reordena la matriz de distancias y se forma un nuevo cluster con el valor más bajo, y así secuencialmente hasta reconstruir el árbol final. El método de unión de vecino comienza con un árbol inicial en forma de estrella que no tiene ramas internas y construye el árbol definitivo localizando secuencialmente “pares de vecinos”, es decir, secuencias de ADN o aminoácidos conectados por un único nodo interior, hasta hallar el árbol en el que las longitudes de ramas internas sean menores, siendo por tanto el árbol de menor longitud.

Por su parte, los métodos de optimización definen un criterio que permite comparar árboles diferentes, para ello, primero se reconstruyen los árboles en base a un algoritmo y después a través de una función objetiva se decide cuál es el mejor. Además, estos métodos proporcionan un rango de bondad de las diferentes soluciones (Swofford *et al.*, 1996). Los métodos de optimización usan la información retenida en las secuencias de ADN de forma más eficiente que los métodos algorítmicos ya que consideran cada posición en el alineamiento. Son computacionalmente más lentos que los métodos de distancia porque tienen que buscar entre varias soluciones posibles, por lo que normalmente se recurre a búsquedas heurísticas más que a búsquedas exhaustivas de los árboles (Swofford *et al.*, 1996). En los métodos de optimización se utilizan aproximaciones heurísticas para generar un árbol inicial usando un algoritmo rápido y después producen reordenamientos locales para mejorar la “puntuación” del árbol inicial, y aunque no garantizan que se encuentre el mejor árbol de entre todos los posibles se incrementa la probabilidad de encontrarlo al aumentar el número de árboles muestreados en el análisis (Yang y Rannala, 2012). Entre los métodos de optimización cabe destacar los métodos de Máxima Parsimonia (MP), Máxima Verosimilitud (ML), e Inferencia Bayesiana (BI). El fundamento teórico y las asunciones previas de cada uno de estos métodos son diferentes y se muestran brevemente a continuación.

El método de Máxima Parsimonia busca el árbol que requiera el menor número de cambios evolutivos (menor longitud) para explicar una matriz de caracteres (Farris, 1970; Fitch, 1971; Swofford *et al.*, 1996). Existen diversos criterios como los de Camin-Sokal, Dollo, Wagner y Fitch para implementar Máxima Parsimonia, basados en probabilidades de cambio que se asumen entre los diferentes estados de caracteres, pero todos son casos especiales del actualmente más común de los métodos de parsimonia que es el método de parsimonia generalizada (Swofford y Olsen, 1990). Este método se basa en la construcción de una matriz que asigna un coste a la transformación de cada estado de carácter en otro y por tanto puede incluir en la misma matriz todos los criterios anteriores basados principalmente en el diferente coste evolutivo que se asigna a la adquisición o reversión de un estado de carácter. Por tanto, incluye en el mismo análisis las ideas de reversibilidad del cambio entre los estados de

caracteres de forma ordenada o desordenada y de irreversibilidad de caracteres (Camin-Sokal, 1965; Kluge y Farris, 1969; Farris, 1970; Fitch, 1971; Farris, 1977). Sin embargo, los métodos de parsimonia ignoran la longitud de ramas, asociada a la probabilidad de cambios evolutivos y asumen independencia de todas las posiciones nucleotídicas, o aminoacídicas en el caso de proteínas (Felsenstein, 2004; Yang, 2006). Además, los análisis de Máxima Parsimonia pueden verse afectados por el efecto de atracción de ramas largas, que da lugar a relaciones espurias cuando el número de caracteres homoplásticos es mayor que los homólogos (Felsenstein, 1978; Bergsten, 2005).

Los métodos de Máxima Verosimilitud e Inferencia Bayesiana son métodos estadísticos basados en modelos de evolución molecular, pero ambos difieren en sus criterios: la Máxima Verosimilitud estima como se ajustan los datos (secuencias de nucleótidos o aminoácidos) a los árboles filogenéticos inferidos de acuerdo a un modelo de evolución propuesto (Swofford *et al.*, 1996; Felsenstein, 2004). La Inferencia Bayesiana estima como se ajustan los árboles filogenéticos a los datos de acuerdo a un modelo evolutivo determinado (Huelsenbeck *et al.*, 2001; Huelsenbeck y Ronquist, 2003). La Máxima Verosimilitud necesita calcular cada árbol posible que pueda ser derivado de la matriz de datos según el modelo de evolución seleccionado, así como su longitud de ramas y las tasas de sustitución entre nucleótidos o aminoácidos (Huelsenbeck y Rannala, 1997; Yang, 2006). Para calcular el valor de probabilidad de un árbol el método de Máxima Verosimilitud primero optimiza las longitudes de rama para calcular el valor de cada árbol candidato y después realiza búsquedas heurísticas en el espacio de árboles para encontrar el árbol que maximiza la función de probabilidad (Yang y Rannala, 2012). La Máxima Verosimilitud puede verse afectada por el efecto de repulsión de grupos hermanos cuando estos se ubican en ramas largas (Siddal, 1998).

La Inferencia Bayesiana como su nombre indica está basada en la estadística bayesiana, con un fundamento teórico distinto a la aproximación de Máxima Verosimilitud. En el caso de la Inferencia Bayesiana, los parámetros del modelo se consideran variables al azar con distribuciones estadísticas más que constantes desconocidas, como ocurre en la Máxima Verosimilitud (Huelsenbeck y Rannala, 2004; Yang y Rannala, 2012). De esta forma, la Inferencia Bayesiana se fundamenta en el teorema de Bayes y expresa que la probabilidad posterior es proporcional a la probabilidad del *prior* multiplicada por la probabilidad del árbol (probabilidad de los datos dado el árbol y el modelo evolutivo). El teorema también introduce la probabilidad marginal de los datos como denominador para asegurar que la suma total de la probabilidad posterior es igual a la unidad. La probabilidad posterior de un árbol es la probabilidad de que dicha topología sea correcta dados los datos y el modelo (Yang y Rannala, 2012). En el análisis de bayes a todos los parámetros del modelo se les asigna un valor de

probabilidad (*prior distribution*) que se combina con los datos para obtener las distribuciones de probabilidad posterior de los parámetros del modelo y de los árboles inferidos (Huelsenbeck y Ronquist, 2001; Ronquist y Huelsenbeck, 2003). Estas probabilidades posteriores no pueden calcularse directamente por su complejidad, por ello se utilizan los algoritmos de cadenas de Monte Carlo (*Markov Chain Monte Carlo* (MCMC); Metropolis *et al.*, 1953; Hastings, 1970) para reducir el tamaño de muestra del espacio de árboles sobre la que estimar las probabilidades posteriores. Las cadenas de Monte Carlo son un algoritmo de simulación que comienzan con un árbol inicial al azar, con longitudes de ramas y parámetros de sustitución también al azar, y van moviéndose de un árbol a otro en función de sus mayores probabilidades posteriores. Sin embargo, para asegurar una correcta aplicación de estas simulaciones, la convergencia entre las cadenas de Markov debe ser evaluada, para lo que se han desarrollado diversas herramientas (Rambaut *et al.*, 2014; Nylander *et al.*, 2008; Ronquist *et al.*, 2012). Esto además permite que a su vez se reduzca el tiempo computacional de la Inferencia Bayesiana respecto a la Máxima Verosimilitud. Otra diferencia entre ambos métodos filogenéticos es que la Inferencia Bayesiana estima simultáneamente el árbol y su incertidumbre (probabilidad posterior), mientras que en el método de Máxima Verosimilitud se necesita implementar un procedimiento adicional para estimar los apoyos estadísticos del árbol, como el *bootstrapping* no paramétrico (Felsenstein, 1985; Hillis y Bull, 1993) para estimar dicha incertidumbre. Sin embargo, se ha criticado que las probabilidades posteriores pueden estar sobrestimadas, especialmente cuando son usados modelos evolutivos demasiado simples (Suzuki *et al.*, 2002; Huelsenbeck y Rannala, 2004; Yang y Rannala, 2005).

Los métodos filogenéticos necesitan incorporar las singularidades de cada proceso evolutivo, para ello se han desarrollado diferentes modelos probabilísticos (modelos evolutivos) que son descripciones probabilísticas del propio proceso de evolución y permiten calcular las probabilidades de cambio entre los distintos estados de caracteres. Dada la gran cantidad de modelos existentes, se necesita un criterio objetivo (*Likelihood ratio test*, *Akaike Information Criterion*, *Bayesian Information Criterion*, etc.) para seleccionar el modelo más apropiado, que es aquel que presenta la mayor probabilidad dados los datos (Akaike, 1973; Schwarz, 1978; Posada y Buckley, 2004; Posada, 2009; Lanfear *et al.*, 2012).

En esta tesis los métodos filogenéticos ML y BI se utilizan ampliamente a lo largo de todos los capítulos y son el pilar sobre el que se sustentan la mayoría de las conclusiones extraídas de los mismos.

5.3. Métodos de Delimitación de especies: puente entre filogenia y genética de poblaciones

La “delimitación de especies” se define como el proceso de identificar la diversidad biológica a nivel de especie, es decir, de definir los límites entre las unidades taxonómicas que se reconocen como especies (Site y Marshall, 2003; Wiens, 2007; Carsten *et al.*, 2013). Sin embargo, la primera limitación a este proceso viene dada por el propio concepto de especie, ya que no existe un consenso entre biólogos sobre una definición que sea válida para todos los campos del conocimiento científico y por ello continúa siendo un debate abierto el concepto teórico de especie (de Queiroz, 2007). Así, el concepto de especie ha sufrido una sustancial transformación desde su definición por Mayr (1942) como “...a group of interbreeding natural populations that is reproductively isolated from other such groups”, es decir, las especies son una consecuencia de la reproducción sexual y el aislamiento reproductivo, y en la actualidad existen más de veinte conceptos diferentes dependiendo de la disciplina científica y del tipo de organismos estudiados (Mayden, 1997; de Queiroz, 2007). Este debate es en parte debido a que las modernas técnicas moleculares han permitido identificar numerosas excepciones al concepto biológico de especies de Mayr. Y es especialmente patente en peces, donde las barreras de aislamiento pre- y poscigóticas son a veces difusas y los casos de hibridación son muy abundantes (ver por ejemplo Freyhof *et al.*, 2005; Gow *et al.*, 2006; Brix y Grossell, 2013).

Son muchos los intentos de definición de especie a lo largo de la literatura científica, ya que la especie es considerada una unidad operativa fundamental en biología (ver Mayden, 1997; de Queiroz, 1998, 2005a,b, 2007 y referencias dentro). Muchas de las definiciones son incompatibles entre sí y pueden derivar en distintas conclusiones y límites difusos entre especies y en la práctica, el concepto de especie está en realidad ligado a los criterios usados para delimitar las especies de forma empírica. Se ha propuesto como posible solución integrar las definiciones de especies en dos criterios: i) un “criterio jerárquico de especie” en el que las diferentes definiciones de especie estarían subordinadas a un único concepto de especie primario, que sería el concepto evolutivo de especie (Mayden, 1997); ii) un “criterio unificado de especie” que pretende retener los elementos comunes de todos los criterios contemporáneos utilizados para definir a la especie y eliminar las partes conflictivas entre todas las definiciones propuestas (de Queiroz, 2007). De acuerdo a sus definiciones, el criterio jerárquico de especies y el criterio unificado de especie formarían parte del mismo marco teórico que consideraría la evolución como un proceso determinante y por ello volvería a cobrar relevancia la definición evolutiva de especie (Wiley 1978; Wiley y Mayden, 2000). De esta forma el concepto evolutivo de especie es adoptado como concepto de especie primario y el conjunto de propiedades biológicas de las especies (aislamiento reproductivo, presencia de caracteres diagnósticos,

monofilia recíproca, diferenciación ecológica, etc.) constituyen criterios operacionales secundarios para delimitar especies (Naomi, 2011). Aún así, el debate teórico continúa abierto, pero se tiende a reconocer a las especies dentro de un marco evolutivo.

Dentro de este marco evolutivo, en la última década las herramientas genéticas se están convirtiendo en un recurso muy útil y muy extendido para definir límites entre especies (ver por ejemplo Pons *et al.*, 2006; Wiens, 2007; Kubatko *et al.*, 2009; O'Meara, 2010; Yang y Rannala, 2010; Fujita *et al.*, 2012; Puillandre *et al.*, 2012; Fujisawa y Barraclough, 2013; Zhang *et al.*, 2013). Así, han surgido nuevas metodologías de delimitación de especies que van un paso más allá de la filogenia e intentan “delimitar” la biodiversidad en unidades discretas introduciendo un algoritmo de coalescencia en la ecuación (Kingman, 1982). De esta forma estos métodos actúan como un puente entre la Filogenia y la Genética de poblaciones, para resolver la incongruencia entre los distintos loci (Maddison y Knowles, 2006; Knowles y Carstens, 2007; Liu *et al.*, 2007; Kubatko *et al.*, 2009; Heled y Drummond, 2010; Fujita *et al.*, 2012). En los modelos de delimitación de especies se incluyen parámetros de filogenia, patrones y tiempo de diversificación de linajes, y parámetros de genética de poblaciones, como son el flujo génico y el tamaño poblacional (Carsten *et al.*, 2013). Sin embargo, estas metodologías aparecidas en los últimos años son muy dispares y sólo suelen ser precisas cuando se utilizan para delimitar especies con unos pocos linajes y una cantidad moderada de *loci*. Además muchas veces es complicado comparar los resultados obtenidos por distintas aproximaciones y por ello se ha propuesto que lo más adecuado es establecer unos intervalos de confianza que sean congruentes entre las distintas metodologías (O'Meara *et al.*, 2010; Yang y Rannala, 2010; Ence y Carstens, 2011; Carstens *et al.*, 2013).

La identificación y delimitación de especies puede ser difícil si sólo se tienen en cuenta caracteres morfológicos, ya que estos caracteres a veces pueden ser confusos por su variabilidad clinal o a lo largo del desarrollo, entre sexos etc. Por ello las metodologías exploratorias basadas en secuencias de ADN como fuente de información primaria para establecer hipótesis de especies en base a la asignación *a priori* de cada individuo a un determinado grupo, que posteriormente deberán ser testadas, se han convertido en una herramienta útil para la sistemática o la taxonomía integrativa (Fujita *et al.*, 2012; Derkarabetian y Hedin, 2014). Las metodologías de delimitación de especies también han demostrado ser útiles para identificar linajes crípticos, en los que los caracteres morfológicos no proporcionan una resolución taxonómica adecuada (Singh *et al.*, 2015; Blair y Bryson, 2017). Algunos de los métodos de delimitación de especies han sido criticados por tener que establecer *a priori* los grupos y se han propuesto metodologías distintas sin necesidad de una asignación previa de grupos para tratar de conferir la mayor objetividad posible a los resultados (Ence y Carsten,

2011; Toprak *et al.*, 2016). Sin embargo, el asignar las muestras a grupos previamente establecidos ha sido considerado un paso clave para obtener resultados fiables (O'Meara, 2010; Carstens *et al.*, 2013). La fiabilidad de las hipótesis de delimitación de especies se refuerza con el uso de aproximaciones multilocus, ya que la exploración de diferentes partes del genoma no ligadas entre sí aumenta su precisión y permite obtener resultados más robustos que si sólo se analiza un único locus, incluso en presencia de procesos como separación incompleta de linajes (Knowles y Carstens, 2007; Yang y Rannala, 2010; Fujisawa *et al.*, 2016).

La mayoría de estas aproximaciones de delimitación de especies multilocus están basadas en el modelo coalescente multiespecie, que generaliza el modelo de deriva genética de Wright-Fisher y se basa en que un modelo de coalescencia puede ser aplicado a los árboles de genes dentro del árbol de especies. Por tanto, permite integrar procesos de coalescencia independientes que ocurren en varias poblaciones conectadas por un árbol evolutivo en el que las ramas constituyen las especies (Rannala y Yang, 2003; Degnan y Rosenberg, 2009; Edwards, 2009). La forma más simple de este modelo asume que el único proceso discordante entre árboles de genes es la separación incompleta de linajes (coalescencias profundas) ya que no existe flujo génico ni introgresión entre las especies. Además, el modelo de coalescencia aplicado en la delimitación de especies tiene otras asunciones como que no existe recombinación ni selección (evolución neutral); que el tamaño poblacional es constante dentro de las poblaciones, aunque no necesariamente entre poblaciones; y que todos los linajes tienen la misma probabilidad de coalescer (Degnan y Rosenberg, 2009). No obstante, esto es una situación teórica ideal ya que en las poblaciones reales algunas de estas premisas no se cumplen (Petit y Excoffier, 2009). Por tanto, los árboles de especies reconcilian las discordancias existentes entre los árboles reconstruidos a partir de los genes individuales, e incluso entre la concatenación de los mismos para generar una supermatriz de nucleótidos, especialmente cuando hay una gran heterogeneidad de árboles de genes (Kubatko y Degnan, 2007). Estos métodos incluyen tanto aproximaciones bayesianas como de máxima verosimilitud así como algunas metodologías basadas en parámetros estadísticos (Edwards *et al.*, 2007; Kubatko *et al.*, 2009; Liu *et al.*, 2009, 2010; Heled y Drummond, 2010; Jewett y Rosenberg, 2012). En los capítulos I y II se emplean metodologías bayesianas para reconstruir árboles de especies, como la propuesta por Heled y Drummond (2010).

5.4. Procesos que pueden oscurecer la filogenia: discordancia entre árboles de genes y árboles de especies

Inferir la historia evolutiva de los organismos debería ser más plausible ahora que existe una gran cantidad de datos moleculares multilocus. Sin embargo, el uso de bases de datos multilocus para reconstruir el árbol consenso de especies de un determinado grupo taxonómico no siempre resulta en las mismas relaciones filogenéticas estimadas por los árboles de genes de forma independiente; de hecho, incluso entre estos árboles de genes observados de forma aislada puede haber incongruencias (Degnan y Rosenberg, 2006; Knowles y Carstens, 2007). En teoría solamente existe un árbol único de especies que refleje la historia evolutiva de un determinado taxón y se ha considerado equivalente a los árboles de genes siempre que fuerzas evolutivas causantes de discordancia entre ambos tipos de árboles (hibridación, duplicación génica, separación incompleta de linajes, transferencia horizontal de genes, etc.) no estén presentes (Maddison, 1997; Maddison y Knowles, 2006). Sin embargo, no deben considerarse equivalentes puesto que están gobernados por premisas conceptuales y procesos de reconstrucción independientes (Knowles y Kubatko, 2010). Son muchos los estudios empíricos que demuestran que este árbol de especies “verdadero” inferido puede ser diferente dependiendo del tipo de marcador molecular utilizado o del análisis realizado, o bien debido a la existencia de procesos evolutivos subyacentes como introgresiones, que pueden estar enmascarando las verdaderas relaciones entre las especies (ver por ejemplo Rokas *et al.*, 2003; Pollard *et al.*, 2006; Song *et al.*, 2012; Salichos y Rokas, 2013).

Una de las discordancias más características entre árboles de genes que dificulta la reconstrucción del árbol filogenético de especies es la discrepancia entre las topologías obtenidas a partir de datos mitocondriales y nucleares para un mismo taxon, y abundantes son los ejemplos en la literatura científica para diversos grupos taxonómicos (ver por ejemplo Sota y Vogler, 2001; Leaché y McGuire, 2006; Linnen y Farrell, 2007; Ting *et al.*, 2008; Spinks y Shaffer, 2009). Dos de los principales factores responsables de la discordancia mitonuclear son entre otros la separación incompleta de linajes y la hibridación. En presencia de estos dos factores se espera que la topología mitocondrial sea incongruente con la nuclear (Maddison, 1997; Funk y Omland, 2003; Linnen y Farrell, 2007). Aunque la diferencia primordial entre la separación incompleta de linajes y la hibridación es que en la separación incompleta de linajes no se espera hallar ningún tipo de patrón biogeográfico mientras que en la introgresión sí es posible encontrarlo, en la práctica, distinguir entre estos dos tipos de procesos es muy complicado (Funk y Omland, 2003; McKay y Zink, 2010; Toews y Brelsford, 2012).

La separación incompleta de linajes, también denominada coalescencia profunda, ocurre

cuando las copias de un gen dentro de una misma especie no encuentran su ancestro común (no coalescen) hasta un tiempo anterior al evento de especiación de forma que los polimorfismos ancestrales pueden mantenerse incluso después del evento cladogenético (Funk y Omland, 2003; Zink y Barraclough, 2008). Esto puede causar que algunos alelos de una especie estén más emparentados con los alelos de otra especie que con los de su propia especie, aún siendo ambas entidades específicas bien definidas. La principal causa de la separación incompleta de linajes es la deriva genética, que provoca la fijación de unos alelos y pérdida de otros por azar. La intensidad a la que se produce la deriva está estrechamente ligada al tamaño poblacional, por tanto, cuanto menor es el tamaño poblacional mayor será su efecto y menos probables serán los procesos de separación incompleta de linajes, ya que hay una menor probabilidad de que se fijen alelos “ancestrales” en distintas poblaciones por azar (Pamilo y Nei, 1988). Por otro lado, por el menor tamaño efectivo del genoma mitocondrial (haploide) respecto del nuclear (diploide), el proceso de separación incompleta de linajes es más probable que ocurra en el genoma nuclear, pues el tiempo de coalescencia es mayor, dando lugar a posibles discordancias mitonucleares (Funk y Omland, 2003). En peces de agua dulce es frecuente observar procesos de separación incompleta de linajes (ver por ejemplo Koblmüller *et al.*, 2010; Tang *et al.*, 2012; Choleva *et al.*, 2014).

La introgresión mitocondrial también puede dar lugar a discordancia mitonuclear, especialmente cuando los niveles de introgresión nuclear son relativamente bajos respecto al mitocondrial (Arnold, 1996; 2007; Funk y Omland, 2003; Avise, 2004; Chan y Levin, 2005; Toews y Brelsford, 2012). Las causas que explican que la introgresión del genoma mitocondrial sea más frecuente que la introgresión de un fragmento del genoma nuclear son varias y entre ellas cabe destacar dos. Una de ellas se fundamenta en las diferencias entre el tamaño de las moléculas de ADN mitocondrial (haploide) y nuclear (diploide), por el mismo proceso ya mencionado de coalescencia profunda, de manera que la separación de linajes en el ADN mitocondrial se completará más rápidamente que en el ADN nuclear (Funk y Omland, 2003; Toews y Brelsford, 2012). Una segunda causa de discordancia mitonuclear viene dada por sistemas de hibridación en los que los machos de una especie se reproducen selectivamente con las hembras de otra especie, de forma que se va reemplazando el ADN mitocondrial de una especie por la otra (Chan y Levin, 2005). La hibridación, y en consecuencia la introgresión genética, es un fenómeno común en peces de agua dulce, en particular en los ciprínidos, incluso entre géneros distintos (ver por ejemplo Scribner *et al.*, 2000; Freyhof *et al.*, 2005; Demandt y Bergeck, 2009). De hecho, la introgresión mitocondrial está ampliamente distribuida a lo largo del árbol de la vida de los peces de agua dulce (ver por ejemplo Tang *et al.*, 2005, 2012; Šlechtová *et al.*, 2007; Šanda *et al.*, 2008; Perdices *et al.*, 2008, 2016). Sin embargo, en muchos grupos de aves y en algunos insectos se ha observado que la introgresión nuclear es más frecuente que la mitocondrial (ver

revisión en Petit y Excoffier, 2009). Es posible diferenciar entre introgresión mitocondrial reciente, cuando los linajes mitocondriales de la especie donadora y receptora no han divergido completamente, aunque este patrón también es frecuente si no se interrumpe el flujo génico (Hey y Nielsen, 2004). Por otro lado, es posible inferir una introgresión genética antigua cuando los linajes mitocondriales han divergido completamente, originando una captura mitocondrial si se interrumpe el flujo génico entre ambas especies y pasa el suficiente tiempo evolutivo para que se separen ambos linajes mitocondriales (Gómez-Zurita y Vogler, 2006; Linnen y Farrell, 2007). Un ejemplo de captura mitocondrial y su contexto evolutivo se expone en el Capítulo V de la presente tesis.

Las discordancias mitonucleares también pueden proceder de presiones de selección diferente sobre los distintos tipos de genomas (mitocondrial y nuclear) o de asimetrías demográficas en procesos de hibridación en los que los tamaños poblacionales de la especie invasora y la especie invadida difieren, o en los que haya una dispersión sesgada hacia uno de los sexos y posterior hibridación con otras poblaciones (Toews y Brelsford, 2012).

5.5. *Estimas de tiempos de divergencia y Biogeografía histórica*

Desde la formulación de la hipótesis de reloj molecular en los años 60, que fue enunciada y testada por Zuckerkandl y Pauling (1962, 1965), y posterior desarrollo de la teoría de evolución neutral de Kimura (1968), la escala temporal ha sido introducida en los estudios evolutivos y la datación molecular de los eventos filogenéticos y biogeográficos ha ganado un interés creciente. La inclusión de esta datación ha modificado enormemente la forma en que se interpretan y exponen los resultados y conclusiones de los estudios evolutivos, ligando procesos como la evolución de un carácter en particular o la aparición de una innovación biológica o ecológica a una escala temporal evolutiva (Forest, 2009). Tras estos trabajos iniciales, Kimura (1968) confirió un fundamento teórico a la idea de reloj molecular postulando su teoría de la evolución neutral, en la que afirmaba que la mayoría de las mutaciones que tienen lugar en el ADN no tienen ningún efecto a nivel molecular (son neutras) y que la aparición de las mismas se produce a un ritmo constante en todos los organismos vivos. A grandes rasgos, el reloj molecular supone que los cambios evolutivos en los fragmentos de ADN (o proteínas) homólogos se acumulan de forma constante a lo largo del tiempo y entre diferentes linajes. Por tanto, analizando estos cambios dentro de un contexto temporal conocido, por ejemplo el registro fósil, es posible estimar el tiempo que ha transcurrido desde la divergencia de dos linajes distintos a partir de su antepasado común.

De estas premisas basadas en que las tasas evolutivas son constantes surge el reloj molecular estricto, utilizado frecuentemente en los primeros trabajos de estimas en los tiempos de divergencia (ver por ejemplo en peces continentales Zardoya y Doadrio, 1999; Machordom y Doadrio, 2001a,b; Doadrio y Perdices, 2001, 2005; Perdices *et al.*, 2002). Sin embargo, se fue acumulando la evidencia de desajustes de este reloj molecular constante debido a la heterogeneidad de tasas entre los distintos organismos como consecuencia de diversos factores intrínsecos a los propios organismos y especie-específicos (tamaño corporal, tasa metabólica, tiempo de generación, tamaño poblacional, presión de selección, etc; Britten, 1996; Rutschmann, 2006; Estabrook *et al.*, 2007; pero ver Whittle y Johnston, 2003). De manera que han surgido nuevas metodologías que intentan corregir dicha heterogeneidad de tasas. Un primer método consiste en eliminar ramas o considerar varias categorías de tasas (Rutschmann, 2006). También se puede integrar la heterogeneidad de tasas en el proceso de datación, usando diferentes modelos de tasas evolutivas o relojes moleculares relajados (Thorne *et al.*, 1998; Sanderson, 1997, 2002; Drummond y Rambaut, 2007). Algunos de los procedimientos que relajan la asunción de reloj molecular asumen que las tasas evolutivas están autocorrelacionadas, es decir, que las tasas de sustitución en los linajes descendientes han sido heredadas de los linajes ancestrales (Thorne *et al.*, 1998; Sanderson, 1997, 2002). En contraposición, otros suponen que estas tasas no están autocorrelacionadas y permiten incorporar su incertidumbre al análisis de datación (Drummond y Rambaut, 2007).

El reloj molecular supuso un gran avance en la Biología Evolutiva porque permite datar los acontecimientos evolutivos cuando el registro fósil está ausente o es ambiguo. No obstante, la aplicación de un reloj molecular de forma correcta requiere de una calibración adecuada. La forma más fiable de calibrar el reloj molecular es mediante el registro fósil, que sin duda provee la información más fiable para tal fin. Sin embargo, a veces el registro fósil es incompleto o la posición filogenética de algún fósil es incierta, de forma que para su ubicación en el árbol filogenético es necesario identificar sinapomorfías compartidas con los grupos actuales, para situar correctamente el fósil en el árbol filogenético (Magallón, 2004; Forest, 2009). La verdadera edad del fósil es desconocida, ya que cuando este se descubre es porque suele ser abundante en el registro fósil, pero no necesariamente cuando realmente surgió el linaje, por ello debe ser considerado como una edad mínima en constricción del nodo en el procedimiento de calibración, es decir, la edad del clado no puede ser más reciente que la edad del fósil (Benton y Ayala, 2003; Magallón, 2004; Near *et al.*, 2005). El uso de un único punto de calibración fósil en el árbol filogenético puede dar lugar a dataciones fuertemente sesgadas, y por ello múltiples calibraciones fósiles deberían ser usadas, aunque no siempre se dispone de un registro fósil apropiado (Conroy y Van Tuinen, 2003; Graur y Martin, 2004; Reisz y Müller, 2004; Near *et al.*, 2005; Rutschmann *et al.*, 2007).

El registro fósil en el género *Squalius* no es muy extenso, la mayoría de los restos conocidos fueron identificados como pertenecientes al género *Leuciscus* Cuvier, 1816, ya que como se ha expuesto previamente en algunas secciones de esta introducción general, el actual género *Squalius* era considerado un subgénero de *Leuciscus* (Berg, 1949; Banareescu, 1964). Fósiles que pueden ser verdaderamente adscritos al género *Squalius* son unos restos hallados en Grecia datados en el periodo Messiniense (*Leuciscus aff. cephalus*, ~6,5 Ma; Yacimiento Lava 2; Ilg y Böhme, 2003); restos en Ucrania datados en el Tortoniense (*Leuciscus cf. cephalus*, ~7,5-8 Ma; Yacimiento Kubanka 2 cerca de Odesa; Ilg y Böhme, 2003); tres escamas halladas en Alemania datadas en el Pleistoceno Inferior, entre 700,000 y 970,000 años, (*Leuciscus cf. cephalus*; Yacimiento Stuttgart Wilhelma Bärengehege; Ilg y Böhme, 2003) y los dientes faríngeos datados en el Mioceno Medio de una especie extinta en la península ibérica (*Leuciscus antunesi*; 13,5-14,5 Ma; Yacimiento Povoia de Santarem, Portugal; Cabrera y Gaudant, 1985; Ilg y Böhme, 2003). Por otro lado, los restos oligocénicos de *Rutilus antiquus* (Cabrera y Gaudant, 1985) podrían ser adscritos al género *Squalius* en base a su similitud morfológica con *S. alburnoides*. En esta tesis se han utilizado a lo largo de los capítulos presentados los fósiles del Mioceno más antiguos para calibrar las dataciones moleculares.

Eventos geológicos como la separación de continentes por tectónica de placas, el surgimiento de una isla volcánica o el levantamiento de una cordillera, etc. también son una fuente de calibración de relojes moleculares. Estos eventos geológicos deben cumplir dos asunciones: el impacto evolutivo del evento geológico debe ser medible (por ejemplo aislamiento de dos linajes) y la edad del evento geológico debe ser conocida con una precisión aceptable (Ho *et al.*, 2005). Por ejemplo la apertura del Estrecho de Gibraltar o el cierre del istmo de Panamá han sido dos episodios geológicos extensamente utilizados para calibrar dataciones moleculares de vertebrados en general y de peces de agua dulce en particular (ver por ejemplo Bermingham *et al.*, 1997; Harris *et al.*, 2002; Doadrio y Perdices, 2001; Machordom y Doadrio, 2001a,b; Perdices *et al.*, 2002; Doadrio y Carmona, 2003; Doadrio y Perdices, 2005; Woodburne, 2010; Ho *et al.*, 2015). La asignación de una calibración geológica a un nodo determinado asume que la divergencia de linajes en ese nodo es el resultado de la aparición de una barrera biogeográfica, sin embargo, a estos eventos geológicos normalmente se les asigna un valor puntual y son procesos que a veces pueden llegar a extenderse millones de años en el tiempo (Forest, 2009). No obstante, la calibración empleando un evento geológico no es adecuada en estudios con un enfoque primordialmente biogeográfico por problemas de circularidad, invalidando las conclusiones extraídas de los análisis de datación (Forest, 2009; Ho *et al.*, 2015). Algunos estudios integran los dos tipos de calibraciones, fósiles más biogeografía, en las dataciones de tiempos de divergencias entre linajes (ver por ejemplo Picq *et al.*, 2014; Pérez-Rodríguez *et al.*, 2016).

Independientemente del uso de relojes moleculares estrictos o relajados, para algunos grupos taxonómicos se han estimado tasas evolutivas a partir de dataciones moleculares calibradas o bien con el registro fósil o bien con eventos geológicos. Estas tasas pueden a su vez utilizarse para calibrar nuevas dataciones moleculares de grupos similares (ver por ejemplo Linder *et al.*, 2005; Forest, 2009). Para los teleósteos, por ejemplo, el rango de tasas evolutivas conocido para el gen mitocondrial citocromo *b*, ampliamente utilizado en esta tesis, es de 0.4-1.1% por linaje por millón de años (ver por ejemplo Zardoya y Doadrio, 1999; Bowen *et al.*, 2001; Machordom y Doadrio, 2001b; Dowling *et al.*, 2002; Doadrio y Perdices, 2005; Near y Bernard, 2004; Berendzen *et al.*, 2008; Perea *et al.*, 2010).

El desarrollo de las dataciones moleculares fue especialmente valioso en el campo de la Biogeografía Histórica, ya que añade una calibración temporal a la direccionalidad demostrada por los eventos filogenéticos en la topología de un árbol que pueden ser relacionados con eventos biogeográficos históricos. Por este motivo, estimas de tiempos de divergencia precisas y fiables son vitales para entender la historia evolutiva de los organismos y los procesos que la han modelado a lo largo del tiempo. La biogeografía histórica es crucial para entender la historia evolutiva de los seres vivos y como la biodiversidad actual ha sido generada. el objetivo de esta disciplina es inferir los patrones espaciotemporales de distribución de los seres vivos e identificar los procesos o factores causales que han modelado dichos patrones a lo largo del tiempo para alcanzar su configuración actual, de forma que además se puedan encontrar patrones comunes entre distintos taxa (Wiens y Donoghue, 2004; Sanmartín, 2012). La biogeografía histórica ha trabajado tradicionalmente considerando una escala espaciotemporal amplia, es decir, analizando patrones que han evolucionado durante millones de años en grandes grupos taxonómicos (nivel supraespecífico) y en grandes escalas espaciales, sin embargo, también se aplica a escalas intermedias de tiempo (miles de años) y a niveles más regionales, e incluso dentro del nivel taxonómico intraespecífico (Sanmartín, 2012).

La biogeografía histórica es una disciplina muy antigua que comienza cuando los primeros exploradores naturalistas empiezan a darse cuenta de que en regiones muy alejadas como África y América del sur aparecen faunas similares. En sus comienzos esta biogeografía histórica tenía una perspectiva puramente narrativa y dispersalista debido al escaso conocimiento sobre la tectónica de placas, que empieza a desarrollarse a mediados del siglo XX (Darlington, 1938; Simpson, 1953; Nelson y Ladiges, 2001). Pero como consecuencia de la introducción de la tectónica de placas en la biogeografía histórica aparece la denominada por algunos autores "biogeografía basada en patrones" (Ronquist, 2003; Sanmartín 2007). Así surge una nueva escuela biogeográfica cuyo objetivo consistía en encontrar patrones generales entre

áreas de endemismo de múltiples organismos sin tener en cuenta ninguna premisa evolutiva subyacente, sino que los procesos evolutivos eran inferidos posteriormente como una interpretación de los resultados obtenidos. Es por tanto una biogeografía principalmente vicariante que descubre patrones de distribución comunes a las grandes biotas alrededor del mundo y postula las principales barreras que separan estas biotas, basándose en las migraciones de los continentes a través del tiempo. Aunque esta primera escuela fue expuesta por primera vez por Croizat (1958) no es tomada en cuenta hasta que la escuela de Nueva York toma interés por la misma y publican, conjuntamente, un demoledor estudio sobre las teorías estáticas, prevalentes hasta entonces, basadas en los centros de origen y dispersión de las faunas (Croizat, 1958; Croizat *et al.*, 1974; Nelson y Platnick, 1981). Las principales críticas hacia los modelos narrativos y dispersalistas, prevalentes hasta entonces, derivan de la falta de soporte científico, ya que las hipótesis que generan no son falsables y su falta de poder predictivo debido a que sólo pueden construir relatos sobre linajes específicos pero no una teoría general para diferentes organismos.

Posteriormente, y debido a la influencia de la metodología de reconstrucción de caracteres de Hennig (1966) y el principio de parsimonia para el estudio filogenético, así como el mejor conocimiento de la deriva continental y la tectónica de placas, se desarrolla la biogeografía cladista que tiene como resultado modelos vicariantes (Brundin, 1966; Croizat *et al.*, 1974; Platnick y Nelson, 1978). La biogeografía cladista es una biogeografía filogenética cuya principal herramienta son los cladogramas de áreas, en los que cada taxón terminal de una filogenia previamente soportada es sustituido por su área de distribución, de manera que comparando cladogramas de distintos organismos que ocurren en la misma región se puede definir una historia biogeográfica común que puede representarse como un cladograma de áreas general (Brooks, 1985, Brooks *et al.*, 2001; Sanmartín, 2007). La biogeografía cladística es un modelo vicariante, y se basa en que un área de distribución es dividida por la aparición de sucesivas barreras biogeográficas. Cada subdivisión de esta área hipotética daría lugar a un evento de especiación alopátrica, por tanto a partir del cladograma de especies podría seguirse la secuencia de eventos vicariantes (Sanmartín, 2012). De esta manera la biogeografía cladista salva las dos críticas de los modelos dispersalistas, por una parte construye hipótesis que pueden ser refutadas en términos de una filosofía estrictamente hipotético-deductiva y puede explicar patrones comunes de distribución en diversos organismos generando teorías biogeográficas generales. El problema viene determinado cuando se estudian floras y faunas con grandes capacidades de dispersión, en ese momento la biogeografía cladística se convierte en extremadamente reduccionista. De esta forma la biogeografía cladística no tiene en cuenta las dispersiones como un proceso evolutivo generador de patrones biogeográficos y considera que estos taxones habrían alcanzado su amplia distribución antes del primer evento de

vicarianza (Humphries y Parenti, 1999; Sanmartín, 2012). La biogeografía cladística tiene varias críticas, una de ellas es que no tiene en consideración los tiempos de divergencia, imprescindible para comparar patrones biogeográficos congruentes, cuando los grupos taxonómicos coinciden en el tiempo y en el espacio, o “pseudocongruentes”, cuando los grupos taxonómicos coinciden en el espacio pero no en el tiempo (Cunningham y Collins, 1994; Donoghue y Moore, 2003). Pero además no incorpora tres elementos fundamentales en biogeografía como son, la dispersión, la extinción y la duplicación de áreas. Además, la comparación entre cladogramas de áreas tampoco se considera suficiente para testar hipótesis biogeográficas generales ya que no realiza ningún tipo de test de contraste de hipótesis (Simberloff *et al.*, 1981; Sanmartín, 2007).

Surge de esta manera la llamada “biogeografía basada en procesos” para solventar el problema del reduccionismo de los modelos cladistas incorporando un modelo determinístico de costes (Ronquist, 1997, 1998, 2003; Sanmartín y Ronquist, 2002). De esta forma, la biogeografía de procesos reconstruye la biogeografía a través de una matriz de costes teniendo en cuenta la distribución ancestral de los organismos y los distintos procesos biogeográficos (vicarianza, dispersión, extinción y duplicación) que pueden ahora integrarse en el propio análisis de reconstrucción (Sanmartín, 2012). A cada evento o proceso biogeográfico se le asigna un coste de acuerdo a la verosimilitud de que dicho proceso haya ocurrido en el pasado. De esta forma la reconstrucción biogeográfica no se hace *a posteriori* interpretando los cladogramas de áreas sino que los procesos biogeográficos son incorporados dentro del propio análisis. De acuerdo a esta aproximación la reconstrucción biogeográfica se obtendría encajando la filogenia de los organismos en un cladograma de áreas que refleja las relaciones entre las áreas de endemismo y la reconstrucción óptima sería aquella que minimice el coste total de los procesos implicados en el patrón que se esté analizando (Ronquist, 2003; Sanmartín, 2007, 2012). El problema es decidir cual es el coste de los eventos biogeográficos. El método más comúnmente utilizado para ello es utilizar un criterio de optimización basado en la parsimonia y elegir los costes que maximicen las oportunidades de obtener patrones de distribución filogenéticamente conservados, es decir aquellos en los que la distribución no cambia del ancestro a los descendientes (Page, 1995; Ronquist, 2003; Sanmartín, 2007, 2012).

Uno de los métodos de reconstrucción de biogeografía histórica basada en la aproximación teórica de biogeografía de procesos es el análisis de dispersión-vicarianza (DIVA). Este método utiliza matrices de costes en las cuales las extinciones y las dispersiones tienen un mayor coste que la vicarianza y la duplicación. Pero además no se basa como otros métodos en un cladograma de áreas sino en mapear las áreas de distribución a través de la filogenia. La hipótesis nula de este análisis es “especiación alopátrica con vicarianza geográfica”, y pretende reconstruir la historia biogeográfica de un grupo de organismos en ausencia de una hipótesis

general de relaciones entre ellos. Para ello, partiendo de una filogenia de un organismo particular y su área de distribución, el análisis de DIVA puede inferir los estados de los nodos con las distribuciones óptimas ancestrales así como los procesos biogeográficos (vicarianzas, extinciones y las dispersiones) que han originado dichas relaciones (Ronquist, 1997; Sanmartín, 2003, 2007). Los eventos de extinción en DIVA se infieren generalmente después del análisis para explicar las distribuciones ancestrales amplias entre áreas que no están geográficamente adyacentes y las dispersiones a larga distancia se infieren también después del análisis considerando las relaciones entre las áreas (Sanmartín, 2003, 2007; Calviño *et al.*, 2008). El análisis de DIVA no tiene en cuenta la incertidumbre de la propia filogenia porque las áreas ancestrales son reconstruidas a partir de una topología fijada y además, múltiples reconstrucciones de áreas con la misma probabilidad pueden dar lugar a soluciones diferentes, por lo que se han desarrollado aproximaciones bayesianas para incorporar esta incertidumbre a los análisis de reconstrucción biogeográfica (Nylander *et al.*, 2008; Harris y Xiang, 2009; Yu *et al.*, 2010, 2015).

Una crítica que se ha realizado a esta metodología es que si el coste estimado es erróneo entonces también lo es el patrón biogeográfico inferido. Sin embargo, los autores que la defienden argumentan la necesidad de tener en cuenta los procesos, y sus costes asociados, para obtener una reconstrucción biogeográfica precisa (Ronquist, 2003; Sanmartín, 2007). Al igual que en la biogeografía cladística, también es difícil incorporar la dimensión tiempo a la reconstrucción biogeográfica basada en parsimonia, ya que el cladograma de áreas refleja el orden relativo de ramificación (equivalente a la diversificación de los linajes), pero no el grado de divergencia entre los linajes o el tiempo desde la cladogénesis, por lo que si se quiere considerar la variable tiempo es necesario incorporar al análisis cladogramas temporalmente calibrados (Sanmartín, 2012). Otras críticas han surgido en estos métodos sobre la inferencia de procesos no ligados a la especiación como la extinción y la dificultad para explicar sucesos como la desaparición de un linaje completamente de un área ancestral o la dispersión completa de un linaje fuera de su área primitiva, porque no dejan una señal filogenética. Es decir procesos como la extinción pueden borrar la señal de vicariancia y crear patrones de distribución artificialmente incongruentes al no incluir información fósil en el análisis biogeográfico. Una ventaja de la biogeografía basada en procesos es que permite testar la significancia estadística de los patrones inferidos comparando los datos observados con los esperados por azar bajo la hipótesis nula de que las distribuciones de los organismos no son filogenéticamente conservadas (Ronquist, 2003; Sanmartín y Ronquist, 2004; Sanmartín, 2007).

Más recientemente y debido a la incorporación de la información generada por los estudios moleculares ha surgido la biogeografía probabilística o paramétrica (Ree *et al.*, 2005;

Ree y Smith 2008; Ree y Sanmartín, 2009; Sanmartín, 2010) que incorpora las estimas de divergencia temporal entre linajes (a través de relojes moleculares), información sobre cambios ambientales y geográficos, el registro fosil o la tolerancia ecológica. Estos métodos están basados en modelos estadísticos de evolución de áreas (cambio de áreas de distribución desde los ancestros hasta los descendientes) cuyos parámetros son los procesos biogeográficos tales como vicarianza, dispersión o extinción (Biogeografía paramétrica o Biogeografía de modelos; Sanmartín, 2012). De esta forma, la evolución del área se modela como un proceso estocástico que cambia a lo largo de las ramas del árbol filogenético de acuerdo a un modelo probabilístico de cadenas de Markov (Sanmartín, 2012). Una limitación de las aproximaciones paramétricas es que asume que la evolución del área está desacoplada con la diversificación de los linajes (Sanmartín, 2012). Uno de los modelos probabilísticos más utilizado en biogeografía paramétrica es el de Dispersión-Extinción-Cladogénesis (DEC) que permite estimar las tasas de expansión (dispersión) o reducción (extinción) del área utilizando una aproximación de máxima verosimilitud y escenarios de reconstrucción de áreas ancestrales a partir de un árbol filogenético calibrado (Ree *et al.*, 2005; Ree y Smith, 2008). Otro modelo probabilístico es el denominado Biogeografía bayesiana de islas (BIB de sus siglas en inglés; Sanmartín *et al.*, 2008), que utiliza la inferencia bayesiana para estimar valores de capacidad de carga de las áreas y tasas de dispersión entre islas a partir de datos de distribución y filogenias de grupos codistribuidos, es decir, introduce una variable ecológica, aunque también se ha utilizado en biogeografía continental (Sanmartín, 2010, 2012).

En los Capítulos I y II de esta tesis se implementan metodologías de biogeografía paramétrica para inferir los patrones biogeográficos de algunas especies concretas del género, así como los procesos evolutivos responsables de dichos patrones biogeográficos.

5.6. Análisis filogeográficos

La creciente facilidad para obtener datos de secuencias de ADN y el surgimiento de la aproximación filogeográfica ha favorecido que el enfoque biogeográfico pueda ser extendible a niveles subespecíficos, trasladando los estudios filogenéticos a este nivel taxonómico cuando las herramientas filogenéticas que funcionan a nivel supraespecífico no tienen el suficiente poder para reconstruir la historia evolutiva de los organismos. La filogeografía es una escuela biogeográfica que reconstruye los patrones y procesos que influyen en la distribución espacial a un nivel poblacional, o entre especies filogenéticamente muy próximas, y a través de marcadores genéticos (Avise *et al.*, 1987; Avise, 2000). Así, la filogeografía ha servido para descubrir especies crípticas (filogrupos) y la comparación de varias de ellas para reconstruir patrones biogeográficos más generales y comunes a diferentes especies (filogeografía comparada;

Bermingham y Moritz, 1998; Arbogast y Kenagy, 2001).

Para algunos autores los límites entre biogeografía y filogeografía son difusos y solamente existiría una diferencia de escala entre ambas aproximaciones. Estos autores tienden a considerar a la filogeografía como una disciplina que se encarga de estudiar patrones biogeográficos recientes como por ejemplo colonizaciones desde refugios climáticos pleistocénicos o patrones geográficos asociados a procesos ecológicos recientes dentro de una misma especie (ver por ejemplo Rissler *et al.*, 2007; Kearns *et al.*, 2014). Sin embargo la filogeografía tiene una serie de propiedades emergentes que no pueden ser incorporadas por la biogeografía clásica, apoyada por metodologías filogenéticas tradicionales de relaciones bifurcadas (Knowles y Maddison, 2002; Avise, 2009). De esta forma, la filogeografía acepta relaciones genealógicas reticuladas resultado de procesos de recombinación o hibridación, es decir múltiples árboles dentro del mismo árbol filogenético, dando lugar a discordancias entre el árbol de especies y los árboles de genes, de forma que no tiene por qué coincidir la divergencia de los distintos linajes genéticos con la divergencia de las poblaciones o con el evento de especiación como consecuencia de fenómenos de coalescencia profunda (Wysor *et al.*, 2012; Mallet *et al.*, 2015). Esto es especialmente importante cuando se pretende estimar la divergencia temporal de las poblaciones dentro de las especies.

Para realizar análisis filogeográficos generalmente se deben utilizar marcadores citoplásmicos no recombinantes, como es el caso del ADN mitocondrial (o cloroplastos en estudios de plantas). Esto permite inferir las relaciones entre los distintos genotipos de los individuos integrantes de las poblaciones objeto de estudio, es decir las relaciones entre los haplotipos, para inferir un árbol genealógico que posteriormente será interpretado desde una perspectiva geográfica y teniendo en cuenta tanto procesos históricos como contemporáneos (Avise, 2009). De esta manera, la forma tradicional más general de realizar un análisis filogeográfico es construir un árbol genealógico a partir de un alineamiento de haplotipos y después representar los linajes geográficamente, pero para ello deben asumirse una serie de asunciones: i) la separación de los linajes debe ser completa, ii) no debe haber flujo génico entre ellos y iii) los ancestros no están representados en la diversidad actual. Si estos criterios no pueden asumirse las secuencias de haplotipos pueden ordenarse gráficamente en una red, que en combinación con su frecuencia y su distribución geográfica provee unas bases sólidas para inferir la historia evolutiva de los organismos vivos (Posada y Crandall, 2001; Emerson y Hewitt, 2005).

En los últimos años, la teoría de la coalescencia ha supuesto un gran avance a la hora de incorporar en los análisis clásicos filogeográficos los procesos de coalescencia profunda y proporciona un riguroso marco estadístico con el que interpretar los árboles genealógicos, considerando a su vez la historia demográfica de las poblaciones, muy asociada a la estructura y diversidad genética de las mismas (Knowles y Maddison, 2002). La coalescencia es un proceso estocástico que se basa en la hipótesis de que todos los genotipos convergen en un ancestro común. Así, los métodos de filogeografía estadística, basados tanto en máxima verosimilitud como en inferencia bayesiana, utilizan simulaciones de coalescencia para generar y contrastar estadísticamente hipótesis alternativas sobre la coalescencia de los linajes y su historia demográfica, para incorporar fenómenos como cuellos de botella, migración, aislamiento genético, expansiones poblacionales, etc, a los análisis filogeográficos (ver por ejemplo Lerney *et al.*, 2009; Hung *et al.*, 2012; De Maio *et al.*, 2015).

Los análisis filogeográficos han sido comúnmente utilizados en la presente tesis para explorar las relaciones genealógicas de las poblaciones de diferentes especies ibéricas del género *Squalius* (Capítulos I, II y IV) y de la vertiente oriental del mar Adriático (Capítulo III), con el fin de describir los patrones geográficos que muestran estas especies así como los procesos asociados a los mismos

6. Modelos de distribución de especies y genética del paisaje: eslabón entre historia evolutiva de las poblaciones y ecología

Uno de los mayores retos actuales es integrar los procesos recientes (ecológicos) en el análisis de la historia evolutiva de los organismos a pesar de que la relación entre factores ambientales, e historia evolutiva de los mismos es un hecho claro, sin embargo, es difícil demostrar causalidad y efecto. Estos dos enfoques, Biología Evolutiva y Ecología, han estado tradicionalmente separados, sin embargo, una forma de unirlos ha sido la creciente disponibilidad de información de distribución de organismos procedente de colecciones científicas o de bases de datos recopilatorias de esta información de distribución de biodiversidad [Global Biodiversity Information Facility (GBIF; <http://www.gbif.org/>); Fishbase (<http://fishbase.org>); Fauna Europaea (<http://www.faunaeurop.org/>); Tree of life (<http://tolweb.org/>), etc.], así como repositorios de secuencias de ADN, principalmente GenBank (<https://www.ncbi.nlm.nih.gov/genbank/>), y de datos ecológicos [Atlas climático digital de la península ibérica (http://opengis.uab.es/wms/iberia/espanol/es_presentacio.htm); worldclim (www.worldclim.org); etc.], ha permitido tener una gran cantidad de información que, junto con el apoyo proporcionado por herramientas como los Sistemas de Información Geográfica (SIG), ha

favorecido el desarrollo de nuevas disciplinas como la genética del paisaje o el modelaje de nicho ecológico que unifican todos estos diferentes enfoques en el mismo análisis (Manel *et al.*, 2003; Storfer *et al.*, 2007, 2010; Elith y Leathwick, 2009).

Una aproximación muy valiosa para investigar la influencia de los factores ambientales sobre la estructura y la diversidad genética es la genética del paisaje, un área del conocimiento que integra la genética de poblaciones con el análisis de variables ambientales y la estadística espacial con el fin de explicar los patrones genéticos observados usando variables ambientales cuantificadas en lugares geográficos concretos (Manel *et al.*, 2003; Holderegger y Wagner, 2008). Asimismo, la genética del paisaje es una disciplina que ayuda a entender los procesos microevolutivos responsables de la estructura y diversidad genética de los distintos organismos a través de una escala espacial (aislamiento por distancia, patrones de estructura genética al azar o determinados por la presencia de barreras al flujo génico, etc). La elección del marcador genético utilizado para el análisis de genética del paisaje dependerá del tipo de cuestión que se pretenda resolver, por ejemplo, no todos los marcadores moleculares son adecuados para estimar determinados parámetros demográficos o poblacionales o no todos los marcadores son válidos para comprender los mecanismos generadores de los patrones espaciales de estructura y diversidad genética observados (Manel *et al.*, 2003). Ejemplos de estudios de genética del paisaje en peces de agua dulce son numerosos en la literatura científica, especialmente aquellos investigando la influencia de la pérdida de conectividad y fragmentación de poblaciones en regiones de climas mediterráneos (ver por ejemplo Huey *et al.*, 2008; Faulks *et al.*, 2010, 2011).

Paralelamente se ha desarrollado una disciplina en torno a la modelización de la distribución de las especies en función de determinadas variables ambientales, es decir, pretenden modelizar su nicho ecológico potencial (Nix, 1986; Elith y Leathwick, 2009; Miller, 2010). Para ello, estos métodos utilizan la distribución espacial actual de un determinado grupo de organismos junto con información obtenida a partir de una serie de variables ambientales para inferir el nicho potencial de dicho grupo, es decir, el conjunto de condiciones bióticas y abióticas en que una especie puede existir y mantener poblaciones viables (ver concepto de nicho de Hutchinson, 1957). Por tanto se puede estimar la distribución potencial de una especie, es decir, el rango geográfico que debería ocupar de acuerdo a su nicho ecológico potencial en relación a las variables ecológicas estudiadas (por ejemplo nicho climático). Esto no quiere decir que sea el nicho actual, sino el que podría ocupar potencialmente. Los modelos de distribución potencial han sido ampliamente utilizados en los dominios terrestre, acuático y marinos, y también han supuesto indirectamente un acercamiento de disciplinas como la Genética y la Ecología, ya que es posible mapear la diversidad genética a lo largo de la distribución de las especies y analizar los principales factores ambientales implicados en la repartición de esta

diversidad genética a través de modelos de regresión lineal, como se expone en el Capítulo IV de esta tesis para *S. valentinus*.

Así, una forma adecuada para investigar la respuesta evolutiva de los organismos al clima se basa en la construcción de modelos de nicho bioclimático (Nix, 1986). Estos modelos establecen las preferencias climáticas de una especie dada en base a su distribución conocida, sin considerar ningún otro tipo de factor ambiental o biótico, y proveen por tanto un modelo de parámetros climáticos correlacionados con dicha distribución. De acuerdo a la teoría de conservacionismo de nicho, los linajes tienden a conservar su nicho ecológico a través del tiempo, es decir, el conjunto de condiciones ambientales en las que los linajes pueden reproducirse y mantener poblaciones viables (Wiens, 2004; Wiens y Donoghue, 2004), y esto puede extrapolarse a variables bioclimáticas. Por ello, el análisis de modelaje de nicho es una de las aproximaciones que puede constituir un nexo de unión entre factores contemporáneos o ecológicos e historia evolutiva de los organismos.

Se han desarrollado multitud de modelos de distribución de especies a lo largo de las últimas tres décadas, entre los que destacan BIOCLIM (Nix, 1986), GARP (Stockwell y Peters, 1999) o MaxEnt (Phillips *et al.*, 2006). Estos modelos presentan una variedad de aplicaciones en biología y en ecología, como es el caso del análisis de distribución de la biodiversidad, la expansión de especies invasoras, simulaciones del impacto del cambio climático en un tiempo futuro. Algunos de estos modelos trabajan solo con datos de presencia, es decir, no tienen en cuenta las ausencias ya que estas a veces no son conocidas, mientras que otros sí que requieren datos de ausencia. Estas aproximaciones que trabajan con datos de presencia realizan una estima de la presencia real de las especies comparando dichos datos de presencia dentro del contexto geográfico del análisis (área total analizada) o con puntos de pseudoausencia estimados, por ello es imprescindible evitar datos de presencia sesgados, es decir, aquellos que no contemplen todo el área de distribución conocida o que procedan de áreas sobremuestreadas, para obtener modelos de distribución fiables (Elith *et al.*, 2009; Phillips *et al.*, 2009).

Uno de los modelos de distribución de especies más utilizado en los últimos años por su facilidad de uso e interpretación de los datos es MaxEnt, el cual está basado en un algoritmo de máxima entropía y en datos de solo presencia, es decir, sin tener en cuenta para el análisis las ausencias (Phillips *et al.*, 2004, 2006; Phillips y Dudík, 2008). En MaxEnt la distribución real de la especie es representada como una función de probabilidad (π) sobre el conjunto de puntos de muestreo (X) del área estudiada, de forma que la función de probabilidad asigna un valor no negativo a cada punto de muestreo (x) y el conjunto de valores $\pi(x)$ suman la unidad. Esta

función de probabilidad a su vez debe cumplir unas constricciones derivadas de los valores de las variables ambientales en los puntos de muestreo, de forma que la media de la variable ambiental estimada por el modelo debe ser próxima al valor empírico observado en los puntos de muestreo, y de todas las distribuciones de probabilidad que satisfacen las constricciones impuestas por las variables se elige la que presente la máxima entropía, es decir la que presenta menores constricciones (Phillips *et al.*, 2006; Phillips y Dudík, 2008).

Estudiar la respuesta evolutiva a cambios históricos climáticos de las especies de agua dulce ayuda a comprender la tendencia futura de estas especies, de vital importancia teniendo en cuenta la creciente tendencia actual a la aridificación de los ecosistemas mediterráneos (Seager *et al.*, 2014). En el caso de la presente tesis se estudia la respuesta evolutiva de la especie *S. valentinus*, endemismo ibérico del levante español, una región donde la estacionalidad mediterránea puede ser extrema (Lázaro *et al.*, 2001), al surgimiento del clima mediterráneo, como se muestra en el Capítulo IV de la misma. Para la obtención de datos de presencia de especies, en la actualidad existen algunas bases de datos a nivel global que recogen información sobre la distribución de las especies (GBIF, etc), aunque el conocimiento de los expertos (bibliografía científica) y las colecciones de museos o herbarios son unas de las fuentes más fiables de datos de distribución de especies (Graham *et al.*, 2004; Elith *et al.*, 2009) y es la que se ha considerado principalmente en esta tesis para realizar los análisis de modelaje de nicho, mostrados en el Capítulo IV.

REFERENCIAS

- Adamson KR, Woodward JC, PD Hughes. 2014. Glaciers and rivers: Pleistocene uncoupling in a Mediterranean mountain karst. *Quaternary Science Reviews*: 94: 28-43.
- Allendorf FW. 1983. Isolation, gene flow, and genetic differentiation among populations. *Genetics and Conservation*. 18(3): 51-65.
- Alley RB, Clark PU, Huybrechts P, Joughin I. 2005. Ice-sheet and sea-level changes. *Science*, 310(5747): 456-460.
- Almada VC, Sousa-Santos C. 2010. Comparisons of the genetic structure of *Squalius* populations (Teleostei, Cyprinidae) from rivers with contrasting histories, drainage areas and climatic conditions based on two molecular markers. *Molecular Phylogenetics and Evolution*, 57(2):924-31.
- Alves MJ, Coelho MM, Collares-Pereira MJ. 2001. Evolution in action through hybridisation and polyploidy in an Iberian freshwater fish: a genetic review. *Genetica*, 111: 375-385.
- Amataj S, Anovski T, Benischke R, Eftimi R, Gorcy LL, Kola L, Leontiadis I, Micevski E, Stamos A, Zoto J. 2007. Tracer methods used to verify the hypothesis of Civijic about the underground connection between Prespa and Ohrid Lake. *Environmental Geology*, 51: 749-753.
- Antonelli A, Sanmartín I. 2011. Mass extinction, gradual cooling, or rapid radiation? Reconstructing the spatiotemporal evolution of the ancient angiosperm genus *Hedyosmum* (Chloranthaceae) using empirical and simulated approaches. *Systematic Biology*, 60(5): 596-615.
- Arai R. 1982. A chromosome study on two cyprinid fishes *Acrossocheilus labiatus* and *Pseudorasbora pumila*, with notes on Eurasian cyprinids and their karyotypes. *Bull. Nat. Sci. Mus. Tokyo Ser. A*, 62: 727-730.
- Arai R y Kato K. 2003. Gross morphology and evolution of the lateral line system and infraorbital bones in bitterlings (Cyprinidae, Acheilognathinae), with an overview of the lateral system in the family Cyprinidae. *Bulletin of the Natural Science Museum. Series A, Zoology*, 40: 1-42.
- Arbogast BS, Kenagy GJ. 2001. Comparative phylogeography as an integrative approach to historical biogeography. *Journal of Biogeography*, 28: 819-825.
- Arnold ML. 1997. *Natural hybridization and evolution*. Oxford University Press. New York. 230 pp.
- Arnold ML. 2006. Evolution through genetic exchange.
- Avice JC. 2000. Phylogeography: the history and formation of species. Cambridge, Mass. Harvard University Press.
- Avice JC. 1998. The history and preview of phylogeography: concepts and applications. *Molecular Ecology*, 7: 371-380.
- Avice JC. 2009. Phylogeography: retrospect and prospect. *Journal of Biogeography*, 36: 3-15.
- Avice JC, Arnold J, Martin Ball R, Bermingham E, Lamb T, Neigel JE, Reeb CA, Saunders NC. 1987. Intraspecific Phylogeography: The mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics*, 18: 489-522.
- Baird IG, Inthaphaisy V, Kisouvannalath P, Phylavanh B, Mounsouphom B. 1999. The fishes of the southern Lao. Lao Community Fisheries and Dolphin Protection Project. Ministry of Agriculture and Forestry, Lao PDR, 161 pp.
- Baker BH, Williams LA, Miller JA, Fitch. 1971. Sequence and geochronology of the Kenia rift volcanics. *Tectonophysics*, 11: 191-215.
- Banarescu P. 1964. *Fauna Republicii Populare Romine: Pisces-Osteichthyes* 13. Bucuresti: AERPR.
- Banarescu P. 1990. Zoogeography of fresh waters. Vol 1. *General distribution and dispersal of freshwater animals*. AULA-VERLAG. Wiesbaden.
- Ball A, Armstrong KF. 2006. DNA barcodes for insect pest identification: a test case with tussock moth (Lepidoptera: Lymantriidae). *Canadian Journal of Forest Research*, 36(2): 337-350.
- Ballard JWO, Rand DM. 2005. The population biology of mitochondrial DNA and its phylogenetic implications. *Annual*

- Review of Ecology and Evolutionary Systematics*, 36: 621-642.
- Bartol M, Pavšič J, Dovnikar M, Bernasconi SN. 2008. Unusual *Braarudosphaera bigelowii* and *Micrantholithus vesper* enrichment in the Early Miocene deiments from the Slovenian Corridor, a seaway linking the Central Parathetys and the Mediterranean. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 267(1-2): 77-88.
- Benton MJ, Ayala FJ. 2003. Dating the tree of life. *Science*, 300: 1698-1700.
- Berendzen PB, Gambleš T, Simons AM. 2008. Phylogeography of the bigeye chub *Hybopsis amblops* (Teleostei: Cypriniformes): early Pleistocene diversification and post-glacial range expansion. *Journal of Fish Biology*, 73: 2021-2039.
- Berg LS. 1949. *Fishes of Fresh Waters of the USSR and Adjacent Countries*. AN SSSR. Moscow, Pt. 2, pp: 469-925 (In Russian).
- Bermingham E, McCafferty SS, Martin AP. 1997. Fish Biogeography and Molecular Clocks: Perspectives from the Panamanian Isthmus. En: Kocher TD, Stepien CA (eds.). *Molecular Systematics of Fishes*. Academic Press. USA., pp: 113-128.
- Bermingham E, Moritz C. 1998. Comparative Phylogeography: concepts and applications. *Molecular Ecology*, 7: 367-369.
- Bianco PG. 1983. *Leuciscus lucumonis* n. sp. from Italy (Pisces: Cyprinidae). *Senckenbergiana Biologica*, 68: 301-309.
- Bianco PG. 1995. Factors affecting the distribution of freshwater fishes especially in Italy. *Cybium*, 19(3): 241-259.
- Bianco PG. 2013. An update on the status of native and exotic freshwater fishes of Italy. *Applied Ichthyology*, 30: 62-77.
- Bianco PG, Ahnelt H, Economidis PS. 1996. The freshwater fishes from eastern, and large Mediterranean islands with comments on their safety status. *Acta Universitatis Carolinae, Biologica*, 40: 45-60.
- Bianco PG, Ketmaier V. 2014. A revision of the *Rutilus* complex from Mediterranean Europe with description of a new genus, *Samarutilus*, and a new species, *Rutilus stoumboudae* (Teleostei, Cyprinidae). *Zootaxa*, 3841(3): 379-402.
- Bianco PG, Nordlie F. 2008. The salinity tolerance of *Pseudophoxinus stymphalicus* (Cyprinidae) and *Valencia letourneuxi* (Valenciidae) from western Greece suggest a revision of the ecological categories of freshwater fishes. *Italian Journal of Zoology*, 75(3): 285-293.
- Blair C, Bryson RW Jr. 2017. Cryptic diversity and discordance in single-locus species delimitation methods within horned lizards (Phrynosomatidae: *Phrynosoma*). *Molecular Ecology Resources*, DOI: 10.1111/1755-0998.12658.
- Blanc P-L. 2002. The opening of the Plio-Quaternary Gibraltar Strait: assessing the size of a cataclysm. *Geodinamica Acta*, 15(5-6): 303-317.
- Blondel J, Aronson J. 1999. *Biology and wildlife of the Mediterranean region*. Oxford University Press. New York. 328 pp.
- Böhme M. 2000. Die Cypriniden (Teleostei, Cypriniformes) des oberoligozänen Maares von Enspel nebst Bemerkungen zur Phylogenie und Biogeographie der Phoxininae. *Paleontol. Z.*, 74(1-2): 99-112.
- Bonada N, Dolédec S, Statzner B. 2007. Taxonomic and biological trait differences of stream macroinvertebrate communities between Mediterranean and temperate regions: implication for future climatic scenarios. *Global Change Biology*, 13(8): 1658-1671.
- Bonada N, Resh VH. 2013. Mediterranean-climate streams and rivers: geographically separated but ecologically comparable freshwater systems. *Hydrobiologia*, 719: 1-29.
- Bonaparte CL. 1846. Catalogo metodico dei pesci Europei. *Atti della Societa Italiana di Scienze Naturali*, 1845: 1-96.
- Bogutskaya NG. 1990a. Morphological principles of systematics of Cyprinid fishes from the subfamily of minnows (Leuciscinae, Cyprinidae). *Vopr. Ikhtiologii*, 30: 920-933. (En Ruso).

- Bogutskaya NG. 1990b. Patterns of variation in some characters of Altai osmans in connection with diagnostic species of the genus *Oreoleuciscus* (Cyprinidae). *Tr. Zool. Inst.*, 222: 110-132.
- Bogutskaya NG. 1990c. Morphological fundamentals in classification of the subfamily Leuciscinae (Cyprinidae). Communication 1. *Journal of Ichthyology*, 30(3): 63-77.
- Bogutskaya NG. 1991. The morphological basis for the classification of cyprinid fishes (Leuciscinae, Cyprinidae) Communication 2. *Journal of Ichthyology*, 31(1): 66-82.
- Bogutskaya NG. 1994. A description of *Leuciscus lepidus* (Heckel, 1843) with comments on *Leuciscus* and leuciscinae-aspinine relationships (Pisces, Cyprinidae). *Ann. Naturhist. Mus. Wien.*, 96B: 599-620.
- Bogutskaya NG. 1996. Contribution to the knowledge of leuciscine fishes of Asia Minor. Part 1. Morphology and taxonomic relationships of *Leuciscus borysthenticus* (Kessler), *Leuciscus smyrnaeus* Boulenger and *Ladigesocypris ghigii* (Gianferrari) (Cyprinidae, Pisces). *Publicación Especial del Instituto Español de Oceanografía*, 21: 25-44.
- Bogutskaya NG. 1997. Contribution to the knowledge of leuciscine fishes of Asia Minor. Part 2. An annotated checklist of leuciscine fishes (Leuciscinae, Cyprinidae) of Turkey with descriptions of a new species and two subspecies. *Mitt. Hamb. Zool. Mus. Inst.*, 94: 161-186.
- Bogutskaya NG. 2002. *Petroleuciscus*, a new genus for the *Leuciscus borysthenticus* species group (Teleostei: Cyprinidae). *Zoosystematics Rossica*, 11: 235-237.
- Bogutskaya NG, Küçük F, Atalay MA. 2006. A description of three new species of the genus *Pseudophoxinus* from Turkey (Teleostei: Cyprinidae: Leuciscinae). *Zoosystematica Rossica*, 15(2): 335-341.
- Böhlen J, Perdices A, Doadrio I, Economidisi PS. 2006. Vicariance, colonisation, and fast local speciation in Asia Minor and the Balkans as revealed from the phylogeny of spined loaches (Osteichthyes, Cobitidae). *Molecular Phylogenetics and Evolution*, 39(2): 552-561.
- Böhme M. 2003. The Miocene Climatic Optimum: evidence from the ectothermic vertebrates of Central Europe. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 195: 389-401.
- Boskovic M, Popovic M, Alilovic N. 2004. Supplement to Skadar lake geogenesis, its inflow and outlet components and background on its regulation activities. En: BALWOIS Conference Proceedings, Ohrid, Macedonia. 25-29 May 2004, pp: 1-9.
- Bostanci D y Polat N. 2009. Age determination and some population characteristics of chub (*Squalius cephalus* L., 1758) in the Çamlidere dam lake (Ankara, Turkey). *Turkish Journal of Science and Technology*, 4(1): 25-30.
- Bowen BW, Bass AL, Rocha LA, Grant WS, Roberstson DR. 2001. Phylogeography of the trumpETFishes (*Aulostomus*): ring species complex on a global scale. *Evolution*, 55: 1029-1039.
- Brawand D, Wagner CE, Li YI, Malinsky M, Keller I, Fan S *et al.* 2014. The genomic substrate for adaptive radiation in African cichlid fish. *Nature*, 513: 375-381.
- Briggs JC. 1995. *Global biogeography*, illustrated ed. Elsevier Science, Arnoldsville.
- Briggs JP. 2002. The zebrafish: a new model organism for integrative physiology. *American Journal of Physiology – Regulatory, Integrative and Comparative Physiology*, 282(1): R3-R9.
- Brikiatis L. 2014. The De Geer, Thulean and Beringia routes: key concepts for understanding early Cenozoic biogeography. *Journal of Biogeography*, 41(6): 1036-1054.
- Briolay J, Galtier N, Brito RM y Bouvet Y. 1998. Molecular phylogeny of Cyprinidae inferred from cytochrome *b* DNA sequences. *Molecular Phylogenetics and Evolution*, 9: 100-108.
- Brito RM, Briolay J, Galtier N, Bouvet N, Coelho MM. 1997. Phylogenetic relationships within genus *Leuciscus* (Pisces: Cyprinidae) in Portuguese freshwaters, based on mitochondrial DNA cythochrome *b* sequences. *Molecular Phylogenetics and Evolution*, 8: 435-442.
- Britten RJ. Rates of DNA-sequence evolution differ between taxonomic groups. *Science*, 231: 1393-1398.
- Brix KV, Grossel M. 2013. Evaluation of pre- and post-zygotic mating barriers, hybrid fitness and phylogenetic

- relationships between *Cyprinodon variegatus variegatus* and *Cyprinodon variegatus hubbsi* (Cyprinodontiformes, Teleostei). *Journal of Evolutionary Biology*, 26(4): 854-866.
- Brusatte SL, Benton MJ, Ruta M, Lloyd GT. 2011. The first 50Myr of dinosaur evolution: macroevolutionary pattern and morphological disparity. *Biology Letters*, 4(6): 733-736.
- Bufalino AP, Mayden RL. 2009. Phylogenetic evaluation of North American Leuciscidae (Actinopterygii, Cyprinoidea) as inferred from analyses of mitochondrial and nuclear DNA sequences. *Systematics and Biodiversity*, 8(4): 493-505.
- Bufalino AP, Mayden RL. 2010a. Phylogenetic relationships of North American phoxinins (Actinopterygii, Cypriniformes: Leuciscidae) as inferred from S7 nuclear DNA sequences. *Molecular Phylogenetics and Evolution*, 55: 143-152.
- Bufalino AP, Mayden RL. 2010b. Phylogenetic relationships of North American phoxinins (Actinopterygii, Cypriniformes: Leuciscidae) based on RAG1 and S7 nuclear DNA data. *Molecular Phylogenetics and Evolution*, 55(1): 274-283.
- Burbrink FT, Pyron RA. 2010. How does ecological opportunity influences rates of speciation, extinction, and morphological diversification in new world ratsnakes (Tribe Lampropeltini)? *Evolution*, 64(4): 934-943.
- Burridge CP, Craw D, Waters JM. 2006. River capture, range expansion, and cladogenesis: the genetic signature of freshwater vicariance. *Evolution*, 60(5): 1038-1049.
- Burridge CP, Craw D, Waters JM. 2007. An empirical test of freshwater vicariance via river capture. *Molecular Ecology*, 16(9): 1883-1895.
- Çağatay MN, Görür N, Flecker R, Sasking M, Tünöglü C, Ellam R, Krigsman W, Vincent S, Dikbas A. 2006. Paratethyan-Mediterranean connectivity in the Sea of Marmara region (NW-Turkey) during the Messinian. *Sedimentary Geology*, 188-189: 171-187.
- Calviño CI, Martínez SG, Downie SR. 2008. The evolutionary history of *Eryngium* (Apiaceae, Saniculoidae): rapid radiations, long distance dispersals and hybridization. *Molecular Phylogenetics and Evolution*, 46: 1129-1150.
- Camin JH, Sokal RR. 1965. A method for deducing branch sequences in phylogeny. *Evolution*, 19: 311-326.
- Carnevale G, Landini W, Sarti G. 2006. Mare versus Lago-Mare: marine fishes and the Mediterranean environment at the end of the Messinian Salinity Crisis. *Journal of the Geological Society*, 163: 75-80.
- Carstens BC, Pelletier TA, Reid NM, Satler JD. 2013. How to fail at species delimitation. *Molecular Ecology*, 22: 4369-4383.
- Cavender TM. 1991. The fossil record of Cyprinidae. En: Winfield IJ, Nelson JS (eds.). *Cyprinid fishes: systematics, biology and exploitation*. Chapman & Hall, Fish and Fisheries. Ser. 3. Londres.
- Cavender TM, Coburn MM. 1992. Phylogenetic relationships of North American Cyprinidae. En: Mayden RL (ed.). *Systematics historical ecology and North American freshwater fishes*. Stanford University Press. Palo Alto, CA, pp: 293-327.
- Carmona JA, Sanjurjo OI, Doadrio I, Machordom A, Vrijenhoek RC. 1997. Hybridogenetic reproduction and maternal ancestry of polyploidy Iberian fish: the *Tropidophoxinellus alburnoides* complex. *Genetics*, 146: 983-993.
- Chan KMA, Levin A. 2005. Leaky prezygotic isolation and porous genomes: rapid introgression of maternally inherited DNA. *Evolution*, 59(4): 720-729.
- Chen W-J, Miya M, Saitoh K, Mayden RL. 2008. Phylogenetic utility of two existing and four novel nuclear gene loci in reconstructing Tree of Life of ray-finned fishes: The order Cypriniformes (Ostariophysi) as a case study. *Gene*, 423(2): 125-134.
- Chen W-J, Lavoué S, Mayden RL. 2013. Evolutionary origin and early biogeography of Ostariophysi fishes (Ostariophysi: Teleostei). *Evolution*, 67(8): 2218-2239.
- Chen X-L, Yue P-Q, Lin R-D. 1984. Major groups within the family Cyprinidae and their phylogenetic relationships. *Acta Zoologica Sinica*, 9: 424-440.

- Chen X-Y. 1996. Breeding tubercles of *Phoxinus* (Teleostei: Cyprinidae). Morphology, distribution and phylogenetic implications. *Journal of Morphology*, 228(2): 127-144.
- Choleva L, Musilova Z, Kohoutova-Sediva A, Paces J, Rab P, Janko K. 2014. Distinguishing between incomplete lineage sorting and genomic introgressions: complete fixation of allospecific mitochondrial DNA in a sexually reproducing fish (*Cobitis*; Teleostei), despite clonal reproduction of hybrids. *PLoS One*, 9(6): e80641.
- Chu YT. 1935. Comparative studies on the scales and on the pharyngeas and their teeth in Chinese cyprinids, with particular reference to taxonomy and evolution. *Biological Bulletin of St. John's University*. Shanghai.
- Coad B. 1981. *Pseudophoxinus persidis*, a new cyprinid fish from Fars, southern Iran. *Canadian Journal of Zoology*, 59(11): 2058-2063.
- Collares-Pereira MJ. 1983. Les phoxinelles circum-méditerranéennes (avec la description d'*Anaocypris* n. gen.) (Poissons, Cyprinidae). *Cybiu*, 7(3): 1-7.
- Collares-Pereira MJ. 1994. Révision caryologique des barbinés et hypothèses concernant la plesiomorphie possible de l'état polyploïde chez les cyprinidés. *Bulletin Français de la Pêche et de la Pisciculture*, 334: 191-199.
- Collares-Pereira MJ, Almacá C. 1979. *Pseudophoxinus hispanicus* (Steindachner, 1866) (Pisces, Cyprinidae). Cyprinidae nouveau pour le Portugal. *Bulletin du Museum Nationale Histoire Naturelle, Paris*, 4e sér. 1, section A 1, pp: 285-287.
- Comber SCL, Smith C. 2004. Polyploid in fishes: patterns and processes. *Biological Journal of the Linnean Society*, 82: 431-442.
- Conroy CJ, Van Tuinen M. 2003. Extracting time from phylogenies: positive interplay between fossil and genetic data. *Journal of Mammalogy*, 84: 444-455.
- Costedoat C, Gilles A. 2009. Quaternary patterns of freshwater fishes in Europe: comparative Phylogeography and conservative perspective. *The Open Conservation Biology Journal*, 3: 36-48.
- Croizat L. 1958. *Panbiogeography*. Volumes 1, 2a, 2b. Caracas, Venezuela.
- Croizat L, Nelson G, Rosen DE. 1974. Centers of origin and related concepts. *Systematic Zoology*, 31:291-304.
- Crow KD, Wagner GP. 2006. What is the role of genome duplication in the evolution of complexity and diversity?. *Molecular Biology and Evolution*, 23(5): 887-892.
- Cuffey KM, Clow GO, Alley RB, Stuiver M, Waddington ED, Saltas RW. 1995. Large arctic temperate change at the Wisconsin-Holocene transition. *Science*, 270: 455-458.
- Cunha C. 2008. The evolution of *Squalius alburnoides* (Teleostei: Cyprinidae) populations in the northern Iberian Peninsula. A molecular genetics and morphological approach into the past and present dynamics of an hybrid complex. Universidad de Lisboa. Tesis doctoral.
- Cunha C, Bastir M, Coelho MM, Doadrio I. 2009. Body shape evolution among ploidy levels of the *Squalius alburnoides* hybrid complex (Teleostei: Cyprinidae). *Journal of Evolutionary Biology*, 22(4): 718-728.
- Cunha C, Coelho MM, Carmona JA, Doadrio I. 2004. Phylogeographical insights into the origins of the *Squalius alburnoides* complex via multiple hybridization events. *Molecular Ecology*, 13(9): 2807-2817.
- Cunha C, Mesquita N, Dowling TE, Gilles A, Coelho MM. 2002. Phylogenetic relationships of Eurasian and American cyprinids using cytochrome *b* sequences. *Journal of Fish Biology*, 61: 929-944.
- Dayrat B. 2005. Towards integrative taxonomy. *Biological Journal of the Linnean Society*, 85(3): 407-415.
- DeFavari J, Zanella LN, Zanella D, Mrakovcic, Merilä J. 2012. Phylogeography of isolated freshwater three-spined stickleback *Gasterosteus aculeatus* populations in the Adriatic Sea. *Journal of Fish Biology*, 80(1): 61-85.
- Degnan JH, Rosenberg NA. 2006. Discordance of species trees with their most likely gene trees. *PLoS Genetics*, 2859: e68.
- Degnan JH, Rosenberg NA. 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends in Ecology and Evolution*, 24(6): 332-340.
- De la Peña A. 1995. Tertiary fishes from the Iberian continental basins: History and fossil record. *Coloquios de*

- Paleontología*, 47: 25-47.
- De Maio N, Wu C-S, O'Reilly K, Wilson D. 2015. New routes to phylogeography: a Bayesian structured coalescent approximation. *PLoS Genetics*, 11(8): e1005421.
- Demandt MH, Bergek S. 2009. Identification of cyprinid hybrids by using geometric morphometrics and microsatellites. *Journal of Applied Ichthyology*, 25(6): 695-701.
- Denys GPJ, Dettai A, Persat H, Doadrio I, Cruaud C, Keith P. 2013. Status of the Catalan chub *Squalius laietanus* (Actinopterygii, Cyprinidae) in France: input from morphological and molecular data. *Knowledge and Management of Aquatic Ecosystems*, 408, 04.
- de Queiroz K. 2005a. A unified concept of species and its consequences for the future of taxonomy. *Proceeding of the National Academy of Sciences of the United States of America*, 102: 6600-6607.
- de Queiroz K. 2005a. Ernst Mayr and the modern concept of species. *Proceeding of the Californian Academy of Sciences*, 56: 196-215.
- de Queiroz K. 2007. Species concepts and species delimitation. *Systematic Biology*, 56: 879-886.
- Derek E, Briggs G, Fortey RA, Wills M. 1992. Morphological disparity in the Cambrian. *Science*, 256: 5064.
- Derkarabetian S, Hedin M. 2004. Integrative taxonomy and species delimitation in harvestmen: a revision of western North American genus *Sclerobunus* (Opiliones: Laniatores: Travunioidea). *PLoS One*, 9(10): e112001.
- Dermitzakis M. 1990. The evolution of the Aegeis during Late Cenozoic. *Geologica Balcanica*, 20: 3-16.
- Dermatzakis M, Papanikolaou DJ. 1981. Paleogeography and Geodynamics of the Aegean region during the Neogene. *Anales Géologiques de Pays Helléniques*, 4: 245-256.
- Di Castri F, Mooney HA (eds). 2012. *Mediterranean-type ecosystems: origin and structure*. Springer-Verlag. New York.
- Doadrio I. 1988. *Leuciscus carolitertii* n. sp. from the Iberian Peninsula (Pisces, Cyprinidae). *Senckenbergiana Biologica*, 68: 301-309.
- Doadrio I, Bouhadad R, Machordom A. 1998. Genetic differentiation and biogeography in Saharan populations of the genus *Barbus* (Osteichthyes, Cyprinidae). *Folia Zoologica*, 47(Suppl. 1): 7-20.
- Doadrio I, Carmona JA. 1998. Genetic divergence in Greek populations of the genus *Leuciscus* and its evolutionary and biogeographical implications. *Journal of Fish Biology*, 53: 591-613.
- Doadrio I, Carmona JA. 2003. Testing freshwater Lago Mare dispersal theory on the phylogeny relationships of Iberian cyprinid genera, *Chondrostoma* and *Squalius*. *Graellsia*, 59(2-3): 457-473.
- Doadrio I, Carmona JA. 2004. Phylogenetic relationships and biogeography of the genus *Chondrostoma* inferred from mitochondrial DNA sequences. *Molecular Phylogenetics and Evolution*, 33: 802-815.
- Doadrio I, Carmona JA. 2006. Phylogenetic overview of the genus *Squalius* (Actinopterygii, Cyprinidae) in the Iberian Peninsula, with description of two new species. *Cybium*, 30(3): 199-214.
- Doadrio I, Casal-López M, Perea S, Yahyhoui A. 2016. Taxonomy of rheophilic *Luciobarbus* Heckel, 1842 (Actinopterygii, Cyprinidae) from Morocco with the description of two new species. *Graellsia*, 72(1): e039.
- Doadrio I, Perea S, Alonso F. 2007b. A new species of the genus *Squalius* Bonaparte, 1837 (Actinopterygii, Cyprinidae) from the Tagus River Basin (Central Spain). *Graellsia*, 63(1): 89-100.
- Doadrio I, Perea S, Alonso F. 2007b. A new species of the genus *Squalius* Bonaparte, 1837 (Actinopterygii, Cyprinidae) from the Tagus River Basin (Central Spain). *Graellsia*, 63(1): 89-100.
- Dobzhansky T. 1937. *Genetics and the origin of species*. Columbia University Press.
- Domínguez-Domínguez O, Vila M, Pérez-Rodríguez R, Remón N, Doadrio I. 2011. Complex evolutionary history of the Mexican stoneroller *Campostoma ornatum* Girard, 1856 (Actinopterygii, Cyprinidae). *BMC Evolutionary Biology*, 11: 153.
- Domínguez-Villar D, Carrasco RM, Pedraza J, Cheng H, Edwards RL, Willenbring JK. 2013. Early maximum extent of paleoglaciers from Mediterranean mountains during the last glaciation. *Scientific Reports*, 2013(3): 2034.

- Douglas ME. 1993. Analysis of sexual dimorphism in an endangered cyprinid fish (*Gila cypha* Miller) using video technology. *Copeia*, 1993: 334-343
- Dowling TE, Brown WM. 1989. Allozymes, mitochondrial DNA, and levels of phylogenetic resolution among four minnow species (*Notropis*: Cyprinidae). *Systematic Biology*, 38(2): 126-143.
- Dowling TE, Tibbets CA, Minckley WL, Smith GR. 2002. Evolutionary relationships of the Plagopterins (Teleostei: Cyprinidae) from cytochrome *b* sequences. *Copeia*, 2002(3): 665-678.
- Drummond AJ, Rambaut A. 2007. BEAST: Bayesian Evolutionary Analysis by Sampling Trees. *BMC Evolutionary Biology*, 7: 214.
- Dupuis JR, Roe AD, Sperling FAH. 2012. Multi-locus species delimitation in closely related animals and fungi: one marker is not enough. *Molecular Ecology*, 21(18): 4422-4436.
- Dudgeon D, Arthington A, Gessner MO, Kawabata Z-I, Knowler DJ, Lévêque C, Naiman RJ, Prieur-Richard A-H, Soto D, Stiassny LJ, Sullivan C. 2006. *Biological Reviews*, 81(2): 163-182.
- Durand J-D, Bianco PG, Laroche J, Gilles A. 2003. Insight into the origin of endemic Mediterranean ichthyofauna: Phylogeography of *Chondrostoma* genus (Teleostei, Cyprinidae). *Journal of Heredity*, 94(4): 315-328.
- Durand J-D, Persat H, Bouvet Y. 1999. Phylogeography and postglacial dispersion of the chub (*Leuciscus cephalus*) in Europe. *Molecular Ecology*, 8(6): 989-997.
- Durand J-D, Tsigenopoulos CS, Ünlü E, Berrebi P. 2002. Phylogeny and biogeography of the family Cyprinidae in the Middle East inferred from cytochrome *b* DNA: Evolutionary significance of this region. *Molecular Phylogenetic and Evolution*, 22: 91-100.
- Durand JD, Ünlü E, Doadrio I, Pipoyan S, Templeton AR. 2000. Origin, radiation, dispersion and allopatric hybridization in the chub *Leuciscus cephalus*. *Proceedings of the Royal Society of London*, B267: 1687-1697.
- Eastman JT. 1977. The pharyngeal bones and teeth of catostomid fishes. *American Midland Naturalist*, 97: 68-88
- Economidis PS. 1991. Check list of Freshwater Fishes of Greece. Recent status of threats and protection. *Hellenic Society of the Protection of Nature*. Athens.
- Economidis PS. 1995. Endangered freshwater fishes of Greece. *Biological Conservation*, 72: 201-211.
- Edwards AWF, Cavalli-Sforza LL. 1963. A method for cluster analysis. *Preprints of the 5th International Biometrics Conference*. Cambridge. UK.
- Edwards AWF, Cavalli-Sforza LL. 1964. Reconstruction of evolutionary trees. *Phenetic and Phylogenetic Classification*. Systematics Association Publication. Vol. 6. London., pp: 67-76.
- Egger B, Koblmüller S, Sturmbauer C, Sefc KM. 2007. Nuclear and mitochondrial data reveal different evolutionary processes in the Lake Tanganyika cichlid genus *Tropheus*. *BMC Evolutionary Biology*, 7: 137.
- Elith J, Leathwick JR. 2009. Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution and Systematics*, 40: 677-697.
- Elmer KR, Reggio C, Wirth T, Verheyen E, Salzburger W, Meyer A. 2009. Pleistocene dessication in East Africa bottlenecked but did not extirpate the adaptive radiation of Lake Victoria haplochromine cichlid fishes. *Proceedings of the National Academy of Sciences of the United States of America*, 106(2): 13404-13409.
- Elith J, Leathwick JR. 2009. Species distribution models: Ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution and Systematics*, 40: 677-697.
- Emerson BC, Hewitt GM. 2005. Phylogeography. *Current Biology*, 15(10): 367-371.
- Englbrecht CC, Freyjo J, Nolte A, Rassman K, Schlieven U, Tautz D. 2000. Phylogeography of the bullhead *Cottus gobio* (Pisces: Teleostei) suggests a pre-Pleistocene origin of the major central European populations. *Molecular Ecology*, 9(6): 709-722.
- Eschmeyer WN, Fricke R, van der Laan R (eds). *Catalog of Fishes: genera, species, references* (<http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp>). Versión electrónica consultada en Diciembre de 2016 [Esta versión fue editada por B. Eschmeyer].

- Esmaeli HR, Khaefi R, Nejad Z. 2016. Historical review on the taxonomy of *Squalius berak* Heckel, 1843 (Teleostei: Cyprinidae). *FishTaxa*, 1(3): 118-126.
- Estabrook GF, Smith GR, Dowling TE. 2007. Body mass and temperature influence rates of mitochondrial DNA evolution in North America. *Evolution*, 61(5): 1176-1187.
- Ewers RM, Didham RK. 2006. Cofounding factors in the detection of species responses to habitat fragmentation. *Biological Review*, 81: 117-142.
- Excoffier L, Foll M, Petit RJ. 2009. Genetic consequences of range expansions. *Annual Review of Ecology and Systematics*, 40:481-501.
- Faulks LK, Gilligan DM, Beheregaray LB. 2010. Islands of water in a sea of dry land: hydrological regime predicts genetic diversity and dispersal in a widespread fish from Australia's arid zone, the golden perch *Macquaria ambigua*. *Molecular Ecology*, 19(21): 4723-4737.
- Faulks LK, Gilligan DM, Beheregaray LB. 2011. The role of Anthropogenic vs. natural in-stream structures in determining connectivity and genetic diversity in an endangered freshwater fish, Macquarie perch (*Macquaria australasica*). *Evolutionary Applications*, 4: 589-601.
- Felsenstein J. 1985. Confidence limits of phylogenies: an approach using the bootstrap. 39(4): 783-791.
- Filipe AF, Araujo MB, Doadrio I, Angermeier PL, Collares-Pereira MJ. 2009. Biogeography of Iberian freshwater fishes revisited: the roles of historical versus contemporary constraints. *Journal of Biogeography*, 36(11): 2096-2120.
- Filipe AF, Lawrence JE, Bonada N. 2013. Vulnerability of stream biota to climate change in Mediterranean climate regions: a synthesis of ecological responses and conservation challenges. *Hydrobiologia*, 719: (1): 331-351.
- Forest F. 2009. Calibrating the Tree of Life: fossils, molecules and evolutionary timescales. *Annals of Botany*, 104(5): 789-794.
- Fountoulis I, Mariolakos I. 2008. Neotectonic folds in the central-western Peloponnese, Greece. *Z. Dtsch. Ges. Geowiss*, 159(3): 485-494.
- Fountoulis I, Mariolakos I, Ladas I. 2014. Quaternary basin sedimentation and geodynamics in SW Peloponnese (Greece) and late stage uplift of Taygetos Mt. *B. Geofis. Teor. Appl.*, 55(2): 303-324.
- Frankham R. 1996. Conservation genetics. *Annual Review of Genetics*, 29: 305-327.
- Franklin R. 1950. The interpretation of diffuse X-ray diagrams of carbon. *Acta Crystallographica*, 3: 107-121.
- Freyhof J, Lieckfeldt D, Bogutskaya NG, Pitra C, Ludwig A. 2006. Phylogenetic position of the Dalmatian genus *Phoxinellus* and description of the newly proposed genus *Delminichthys* (Teleostei: Cyprinidae). *Molecular Phylogenetics and Evolution*, 38(2): 416-425.
- Freyhof J, Özüluğ M. 2010. *Pseudophoxinus hittitorum*, a new species of spring minnow from central Anatolia (Teleostei, Cyprinidae). *Ichthyological Exploration of Freshwaters*, 21(3): 239-245.
- Freyhof J, Özüluğ M. 2011. Revision of the genus *Squalius* in Western and Central Anatolia, with description of four new species (Teleostei: Cyprinidae). *Ichthyological Exploration of Freshwaters*, 22(2): 107-148.
- Frisch W, Kuhlemann J, Dunkl I, Brügel A. Palinspastic reconstruction and topographic evolution of the Eastern Alps during the late Tertiary extrusion *tectonophysics*, 297: 1-15.
- Fujita MK, Leaché AD, Burbrink ET, McGuire JA y Moritz C. 2012. Coalescent-based species delimitation in an integrative taxonomy. *Trends in Ecology and Evolution*, 27: 480-488.
- Gante HF, Micael J, Oliva-Paterna FJ, Doadrio I, Dowling TE, Alves MJ. 2009. Diversification within glacial refugia: tempo and mode of evolution of the polytypic fish *Barbus sclateri*. *Molecular Ecology*, 18: 3240-3255.
- Gante HF, Santos CD, Alves MJ. 2007. A new species of *Chondrostoma* Agassiz, 1832 (Cypriniformes: Cyprinidae) with sexual dimorphism from the lower river Tejo Basin, Portugal. *Zootaxa*, 1616: 23-35.
- Gasith A, Resh VH. 1999. Streams in Mediterranean climate regions: abiotic influences and biotic responses to predictable seasonal events. *Annual Review of Ecology and Systematics*, 30: 51-81.

- Gaubert P, Denys GP, Oberdorff T. 2009. Genus-level supertree Cyprinidae (Actinopterygii: Cypriniformes), partitioned qualitative clade support and test scenarios of macro-evolutionary scenarios. *Biological Reviews*, 84: 653-689.
- Gaudant J. 1984. Nouvelles recherches sur les cyprinidae (Poisson, Téléostéens) dans l'oligocène des limagnes. *Geobios*, 17: 659-666.
- Geiger MF, Herder F, Monaghan M, Almada V, Barbieri R, Bariche M, Berrebi P, Bohlen J, Casal-Lopez M, Delmastro G, Denys G, Dettai A, Doadrio I, Kalogianni E, Käst H, Kottelat M, Kovacic M, Laporte M, Lorenzoni M, Marcic Z, Özulug M, Perdices A, Perea S, Persat H, Porcellotti S, Puzzi C, Robalo J, Sanda R, Schneider M, Slechtova V, Stumboudi M, Walter S, Freyhof J. 2014. Spatial heterogeneity in the Mediterranean Biodiversity Hotspot affects barcoding accuracy of its freshwater fishes. *Molecular Ecology Resources*, 14: 1210-1221.
- Gianneto D, Carosi A, Ghetti L, Pedicillo G, Pompei L, Lorenzoni M. 2013. Ecological traits of *Squalius lucumonis* (Actinopterygii, Cyprinidae) and main differences in the Tiber River Basin (Italy). *Knowledge and Management of Aquatic Ecosystems*, 409: 04.
- Gilles A, Lecomte G, Faure E, Chappaz R, Brun G. 1998. Mitochondrial phylogeny of the European cyprinids: implications for their systematics, reticulate evolution, and colonization time. *Molecular Phylogenetics and Evolution*, 19: 22-33.
- Gómez A, Lunt DH. 2007. Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula. En: Weiss S, Ferrand N (eds.). *Phylogeography of Southern European Refugia*. Springer, Dordrecht. pp: 155-188.
- Gosline WA. 1978. Unbranched dorsal-fin rays and subfamily classification in the fish family Cyprinidae. *Occasional Papers of the Museum of Zoology, University of Michigan*, 684: 1-21.
- Gousskov A, Vorburger C. 2016. Postglacial recolonizations, watershed crossing and human translocations shape the distribution of chub lineages around the Swiss Alps. *BMC Evolutionary Biology*, 16: 185.
- Gow JL, Peichel CL, Taylor EB. 2006. Contrasting hybridization rates between sympatric three-spined sticklebacks highlight the fragility of reproductive barriers between evolutionarily young species. *Molecular Ecology* 15: 739-752.
- Graham CH, Ferrier S, Huettman F, Moritz C, Peterson AT. 2004. New developments in museum-based informatics and applications in biodiversity analysis. *Trends in Ecology and Evolution*, 19: 497-503.
- Graham CH et al., 2006. Habitat history improves prediction of biodiversity in a rainforest fauna. *Proceedings of the National Academy of Science of the United States of America*, 103: 632-636.
- Graur D, Martin W. 2004. Reading the entrails of chickens: molecular timescales of evolution and the illusion of precision. *Trends in Genetics*, 20: 80-86.
- Groenewald AAVJ. 1958. A revision of the genera *Barbus* and *Varicorhinus* (Pisces: Cyprinidae) in transvaal. Annual Transvaal Museum, 23 (Pt. 23): 263-340.
- Grunwald DJ, Elsen JS. 2002. Headwaters of the zebrafish – emergence of a new model vertebrate. *Nature Review Genetics*, 3: 717-724.
- Guinand B, Durand J-D, Laroche J. 2001. Identifying main evolutionary mechanisms shaping genetic variation of *Leuciscus cephalus* L., 1758 (Cyprinidae) in Western Greece: discordance between methods. *Comptes Rendus de l'Académie des Sciences – Series III – Sciences de la Vie*, 324(11): 1045-1060.
- Hallgrissom B, Jamniczky HA, Young NM, Rolian C, Schmidt-ott U y Marcucio RS. 2012. The generation of variation and the developmental basis for evolutionary novelty. *Journal of Experimental Zoology*, 318(6): 501-571.
- Hänfling B, Brandl R. 1998. Genetic and morphological variation in a common European cyprinid, *Leuciscus cephalus* within and across Central European drainages. *Journal of Fish Biology*, 52(4): 706-715.
- Hänfling B, Brandl R. 2000. Phylogenetics of European cyprinids; insights from allozymes. *Journal of Fish Biology*, 57: 265-276.

- Hänfling B, Dümpelmann C, Bogutskaya NG, Brandl R, Brändle M. 2009. Shallow phylogeographic structuring of *Vimba vimba* across Europe suggests two distinct refugia during the Last Glaciation. *Journal of Fish Biology*, 75(9): 2269-2286.
- Hänfling B, Hellemans B, Volckaert FAM, Carvalho GR: 2002. Late glacial history of the cold-adapted freshwater fish *Cottus gobio*, revealed by microsatellites. *Molecular Ecology*, 11: 1717-1729.
- Hare MP. 2001. Prospects for nuclear gene phylogeography. *Trends in Ecology and Evolution*, 16: 700-706.
- Harris DJ, Carranza S, Arnold EN, Pinho C, Ferrand N. 2002. Complex biogeographical distribution of genetic variation within *Podarcis* wall lizards across the Strait of Gibraltar. *Journal of Biogeography*, 29: 1257-1262.
- Harmon LJ, Schulte JA, Larson A y Losos JB. 2003. Tempo and mode of evolutionary radiation in Iguanian lizards. *Science*, 301 (5635): 961-964.
- Hays JD, Imbrie J, Shackleton NJ. 1976. Variations in the Earth's orbit: the pacemaker of the ice ages. *Science*, 194: 1121-1132.
- Haywood AM, Dowsett HJ, Dolan AM. 2016. Integrating geological archives and climate models for the mid-Pliocene warm period. *Nature Communications*, 7: 10646.
- Hazhauser M, Piller WE, Steininger FF. 2002. Circum-Mediterranean Oligo-Miocene biostratigraphic evolution – the gastropod's point of view. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 183: 103-133.
- Hazhauser M, Kroh A, Mandic O, Piller WE, Göhlich U, Reuter M, Berning B. 2007. Biogeographic responses to geodynamics: A key study all around the Oligo-Miocene Tethyan Seaway. *Zoologischer Anzeiger*, 246: 241-256.
- He SP, Mayden RL, Wang X, Wang W, Tang KL, Chen W-J, Chen Y. 2008. Molecular phylogenetics of the family Cyprinidae (Actinopterygii: Cypriniformes) as evidenced by sequence variation in the first intron of S7 ribosomal protein coding gene: Further evidence from a nuclear gene of the systematic chaos in the family. *Molecular Phylogenetics and Evolution*, 46: 818-829.
- Hebert PDN, Cywinska A, Ball SL y de Waard JR. 2003. Biological identification through barcodes. *Proceedings of the Royal Society B. Biological Sciences*, 270: 313-321.
- Hebert PDN, Gregory T. 2005. The promise of DNA barcoding for taxonomy. *Systematic Biology*, 54: 852-859.
- Heller J. 2007. A historic biogeography of the aquatic fauna of the Levant. *Biological Journal of the Linnean Society*, 92(4): 625-639.
- Hendry AP y Kinnison MT. 2001. An introduction to microevolution: rate, patterns, process. *Genetica* (1-8), 112-113.
- Hrbek T, Seckinger J, Meyer A. 2007. A phylogenetic and biogeographic perspective on the evolution of poeciliid fishes. *Molecular Phylogenetics and Evolution*, 43(3): 986-998.
- Hewitt GM. 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*, 58(3): 247-276.
- Hewitt GM. 2004. Genetic consequences of climatic oscillations in the Quaternary. *Philosophical transactions of the Royal Society B. Biological Sciences*, 359(1442):
- Hillis DM, Bull JJ. 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. *Systematic Biology*, 42(2): 182-192.
- Hinsbergen DJ Jr, Dupont-Nivet G, Nakov R, Ouf K, Panaiotu C. 2008. No significant post-Eocene rotation of the Moesian platform and rhodope (Bulgaria): implications for the kinematic evolution of the Carpathian and Aegean arcs. *Earth Planet Science Letters*, 273: 345-358.
- Ho SYW, Tong J, Foster CSP, Ritchie AM, Lo N, Crisp MD. 2015. Biogeographic calibrations for the molecular clock. *Biology Letters*, 11(9): 20150194.
- Holderegger R, Wagner HH. 2008. Landscape genetics. *BioScience*, 58: 199-207.
- Holling CS. 1973. Resilience and stability of ecological systems. *Annual Review of Ecology and Systematics*, 4: 1-23.
- Howes GJ. 1991. Systematics and biogeography: An overview. En: Winfield IJ y Nelson JS (eds.). *Cyprinid fishes: Systematics biology and exploitation*. Chapman & Hall, London, UK. Pp: 1-3.

- Hughes JM, Schmidt DJ, Fin DS. 2009. Genes in streams: Using DNA to understand the movement of freshwater fauna and their riverine habitat. *BioScience*, 59(7): 573-583.
- Hughes PD, Woodward JC. 2008. Timing of glaciation in the Mediterranean mountains during the last cold stage. *Journal of Quaternary Science*, 23(6-7): 575-588.
- Hughes PD, Woodward JC, Gibbard PL. 2006. Late Pleistocene glaciers and climate in the Mediterranean. *Global and Planetary Science*, 50(1-2): 83-98.
- Hughes PD, Woodward JC, Gibbard PL. 2006. Middle Pleistocene cold stages climates in the Mediterranean: New evidence from the glacial record. *Global and Planetary Science*, 253(1-2): 50-56.
- Huey JA, Baker AM, Hughes JM. 2008. The effect of landscape processes upon gene flow and genetic diversity in an Australian freshwater fish, *Neosilurus hyrtii*. *Freshwater Biology*, 53(7): 1393-1408.
- Hung C-M, Drovetski SV, Zink RM. 2012. Multilocus coalescence analyses support A mtDNA-based phylogeographic history for a widespread palearctic passerine bird, *Sitta europaea*. *Evolution*, 66(9): 2850-2864.
- Hutchinson GE. 1957. Concluding remarks. *Cold Spring Harbor Symp Quantitative Biol.*, 22: 415-427.
- Hyuse T, Van Houdt J, Volckaert FAM. 2004. Paleoclimatic history and vicariant speciation in the "sand goby" group (Gobiidae, Teleostei). *Molecular Phylogenetics and Evolution*, 32: 324-336.
- Imoto JN, Saitoh K, Sasaki T, Yonezawa T, Adachi J, Kartavsev YP, Miya M, Nishida M y Hanzawa N. 2013. Phylogeny and biogeography of highly diverged freshwater fish species (Leuciscinae, Cyprinidae, Teleostei) inferred from mitochondrial genome analysis. *Gene*, 514: 112-124.
- Ito Y, Sakai H, Shekdo S, Jeon S-R. 2002. Genetic differentiation of the northern Far East cyprinids, *Phoxinus* and *Rhynchocypris*. *Fisheries science*, 68(1): 75-78.
- Jablonski D. 2000. Micro- and Macroevolution: scale and hierarchy in evolutionary biology and paleobiology. *Paleobiology*, 26(4): 15-52.
- Jiménez-Moreno G, Fauquette S, Suc J-P. 2010. Miocene to Pliocene vegetation reconstruction and climate estimates in the Iberian Peninsula from pollen data. *Review of Paleobotany and Palynology*, 162(3): 403-415.
- Khaefi R, Esmaeili HR, Sayyadzadeh G, Geiger MF, Freyhof G. 2016. *Squalius namak*, a new chub from Lake Namak basin in Iran (Teleostei: Cyprinidae). *Zootaxa* 4169(1): 145-159.
- Kartavtsev YP, Hanzawa N. 2007. Inferences in Leuciscinae (Pisces, Cyprinidae) phylogeny and taxonomy based on cytochrome b sequences distances and on enzymes loci diversity. *Korean Journal of Genetics*, 29(4).
- Kartavtsev YP, Sviridov VV, Hanzawa N, Sasaki T. 2002. Genetic divergence of far-eastern dace species belonging to the genus *Tribolodon* (Pisces, Cyprinidae) and closely related taxa. *Russian Journal of Genetics*, 38(11): 1285-1297.
- Kearns AM, Joseph L, Toons A, Cook LG. 2014. Australia's arid-adapted butcherbirds experienced range expansions during Pleistocene glacial maxima. *Nature Communications*, 5: 3994.
- Kéfi S, Rietkerk M, Alados CL, Pueyo Y, Papanastasis VP, ElAich A, de Ruiter PC. Spatial vegetation patterns and imminent desertization in Mediterranean arid ecosystems. *Nature*, 449: 213-217.
- Ketmaier V, Bianco PG, Cobolli M, de Matthaes E. 2003. Genetic differentiation and biogeography in southern European populations of the genus *Scardinius* (Pisces, Cyprinidae) based on allozyme data. *Zoologica Scripta*, 32(1): 13-22.
- Ketmaier V, Bianco PG, Cobolli M, Krivokapic M, Caniglia M, de Matthaes E. 2004. Molecular phylogeny of two lineages of Leuciscinae cyprinid (*Telestes* and *Scardinius*) from peri-Mediterranean area based on cytochrome b data. *Molecular phylogenetics and Evolution*, 32: 1061-1071.
- Ketmaier V, Cobolli M, de Matthaes E, Bianco PG. 1998. Allozymic variability and biogeographic relationships in two *Leuciscus* species complexes (Cyprinidae) from southern Europe, with the rehabilitation of the genus *Telestes* Bonaparte. *Italian Journal of Zoology*, 65: 41-48.
- Kimura M. 1968. Evolutionary rate at the molecular level. *Nature*, 217: 624-626.

- Knowles LL, Carstens B. 2007. Delimiting species without monophyletic gene trees. *Systematic Biology*, 56: 887-895.
- Knowles LL, Maddison WP. 2012. Statistical phylogeography. *Molecular Ecology*, 11: 2623-2635.
- Kobmüller S, Egger B, Sturmbauer C, Sefc KM. 2010. Rapid radiation, ancient incomplete lineage sorting and ancient hybridization in the endemic lake Tanganyika cichlid tribe Tropheini. *Molecular Phylogenetics and Evolution*, 55(1): 318-334.
- Kocher TD, Stepien CA. 1997. *Molecular Systematics of Fishes*. Academic Press. USA.
- Kocher TD, Thomas WK, Meyer A, Edwards SV, Pääbo S, Villablanca FX, Wilson AC. 1989. Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proceedings of the National Academy of Sciences of United States of America*, 86(16): 6196-6200.
- Kocher TD, White TJ. 1989. Evolutionary Analysis via PCR. En: Erlich HA (ed.). PCR Technologies. Principles and Applications for DNA Amplification. Palgrave MacMillan, UK., pp: 137-147.
- Koehn RK. 1965. Development and ecological significance of nuptial tubercles of the red shiner, *Notropis lutrensis*. *Copeia*, 1965(4): 462-467.
- Kokkalas S, Xypolias P, Koukouvelas I, Doutsos T. 2006. Postcollisional contractional and extensional deformation in the Aegean region. En: Dilek Y, Pavlides S (eds.). Postcollisional tectonics & magmatism in the Mediterranean region and Asia. *Geol. Soc. Am. S.*, 409: 97-123.
- Kontula T, Väinölä R. 2001. Postglacial colonization of northern Europe by distinct phylogeographic lineages of the bullhead, *Cottus gobio*. *Molecular Ecology*, 10(8): 1983-2002.
- Kottelat M. 1997. *European freshwater fishes*. Biologia (Bratislava), 52(Suppl. 5). 271 pp.
- Kottelat M, Barbieri R. 2004. *Pseudophoxinus laconicus*, a new species of minnow from Peloponnese, Greece, with comments on the west Balkan *Pseudophoxinus* species (Teleostei: Cyprinidae). *Ichthyological Exploration of Freshwaters*, 15(2): 147-160. *Squalius orpheus*, a new species of cyprinid fish from Evros drainage, Greece (Teleostei, Cyprinidae). *Ichthyological Exploration of Freshwaters*, 17(2): 181-186.
- Kottelat M, Economidis PS. 2006.
- Kottelat M, Freyhof J. 2007a. Handbook of the European Freshwater Fishes. Cornol. Switzerland. 646 pp.
- Kottelat M, Freyhof J. 2007b. *Pelagus* a new genus name for the Balkan species *Pseudophoxinus* (Teleostei, Cyprinidae). *Ichthyological Exploration of Freshwaters*, 18(2): 103-108.
- Kotlík P, Bogutskaya NG, Ekmekçi FG. 2004. Circum-Black sea phylogeography of *Barbus* freshwater fishes: divergence in the Pontic glacial refugium. *Molecular Ecology*, 13(1): 87-95.
- Kubatko LS, Carstens BC, Knowles LL. 2009. STEM: species tree estimation using maximum likelihood for gene trees under coalescence. *Bioinformatics*, 25(7): 971-973.
- Kuhlemann J. 2007. Palaeogeographic and paleotopographic of the Swiss and Eastern Alps since the Oligocene. *Global Planet Change*, 58: 224-236.
- Lake PS. 2003. Ecological effects of perturbation by drought in flowing waters. *Freshwater Biology*, 48(7): 1161-1172.
- Lara A, Ponce de León JL, Rodríguez R, Casane D, Côté G, Bernatchez L, García-Machado E. 2009. DNA barcoding of Cuban freshwater fishes: evidence for cryptic species and taxonomic conflicts. *Molecular Ecology Resources*, 10(3): 421-430.
- Lázaro R, Rodrigo FS, Gutiérrez L, Domingo F, Puigdefábregas J. 2001. Analysis of a thirty year rainfall record (1967-1997) from semi-arid SE Spain: a plant ecological perspective. *Journal of Arid Environments*, 48: 373-395.
- Leaché AD, McGuire JA. 2006. Phylogenetic relationships of horned lizards (*Phrynosoma*) based on nuclear and mitochondrial data: evidence for a misleading mitochondrial gene. *Molecular Phylogenetics and Evolution*, 39(3): 628-644.
- Lerney P, Rambaut A, Drummond AJ, Suchard MA. 2009. Bayesian phylogeography find its roots. *PLoS One*, 5(9): 1-5.

e1000520.

- Levin BA, Freyhof J, Lajbner Z, Perea S, Abdoli A, Gaffaroglu M, Ozulug M, Rubenyan HR, Salnikov VB, Doadrio I. 2012. Phylogenetic relationships of the algae scrapping cyprinid genus *Capoeta* (Teleostei, Cyprinidae). *Molecular Phylogenetic and Evolution*, 62(1): 542-549.
- Levy A, Doadrio I, Almada VC. 2009. Historical biogeography of European leuciscins (Cyprinidae): evaluating the Lago Mare dispersal hypothesis. *Journal of Biogeography*, 36(1): 55-65.
- Liao T-Y, Ünlü E, Kullander EO. 2011. Western boundary of the subfamily Danioninae in Asia (Teleostei, Cyprinidae): derived from the systematic position of *Barilius mesopotamicus* based on molecular and morphological data. *Zootaxa*, 2880: 31-40.
- Linder HP, Hardy CR, Rutschmann F. 2005. Taxon sampling effects in molecular clock: An example from the African Restionaceae. *Molecular Phylogenetic and Evolution*, 35(3): 569-582.
- Liu H y Chen Y. 2003. Phylogeny of the East Asian cyprinids inferred from sequences of the mitochondrial DNA control region. *Canadian Journal of Zoology*, 81: 1938-1946.
- Liu H, Tzeng CS y Teng Y. 2002. Sequence variations in the mitochondrial DNA control region and their implications for the phylogeny of the Cypriniformes. *Canadian Journal of Zoology*, 80: 569-581.
- Logez M, Bady P, Melcher A, Pont D. 2013. A continental-scale analysis of fish assemblage functional structure in European Rivers. *Ecography*, 36: 80-91.
- López-Martínez N. 1989. Tendencias en paleogeografía. El future de la biogeografía del pasado. En: Aguirre E (ed.). *Paleobiología*. CSIC. Madrid, pp: 271-296.
- Lorenzoni M, Carosi A, Pedicillo G, Pompei L, Rocchini M. 2011. Reproductive properties of the chub *Squalius squalus* (Bonaparte, 1837) in the Assino creek (Umbria, Italy). *Knowledge and Management of Aquatic Ecosystems*, 403: 09.
- Lossos JB, Glor RE. 2003. Phylogenetic comparative methods and the geography of speciation. *Trends in Ecology and Evolution*, 18(5): 220-227.
- Lowe WH, Allendorf FW. 2010. What can genetics tell us about population connectivity? *Molecular Ecology*, 19: 3038-3051.
- Lowe S, Browne M, Boudjelas S, De Poorter M. 2000. 100 of the World's Worst Invasive Alien Species. A selection from the global invasive species. Published by The Invasive Species Specialist group (ISSG) a specialist group of the Species Survival Commission (SSC) of the World Conservation Union (IUCN). 12pp.
- Lydeard C, Roe K. 1997. The phylogenetic utility of the mitochondrial cytochrome b gene for inferring relationships among Actinopterygian fishes. En: Kocher TD, Stepien CA (eds.). *Molecular Systematics of Fishes*. Academic Press. San Diego. USA. pp: 285-303.
- Machordom A, Doadrio I. 2001a. Evolutionary history and speciation modes in the cyprinid genus *Barbus*. *Proceedings of the Royal Society of London, Series B*, 268: 1297-1306.
- Machordom A, Doadrio I. 2001b. Evidence of a Cenozoic Betic-Kabilian connection based on freshwater fish phylogeography (*Luciobarbus*, Cyprinidae). *Molecular Phylogenetics and Evolution*, 18(2): 252-263.
- Machordom A, Doadrio I, Berrebi P. 1998. Phylogeny and evolution of the genus *Barbus* (Osteichthyes: Cyprinidae) in the Iberian Peninsula as revealed by allozyme electrophoresis. *Journal of Fish Biology*, 47: 211-236.
- Machordom A, Bouhadad R, Doadrio I. 1995. Allozyme variation and evolutionary history of North African populations of the genus *Barbus* (Osteichthyes: Cyprinidae). *Diversity and Distributions*, 4: 217-234.
- Magallón SA. 2004. Dating lineages: molecular and paleontological approaches to the temporal framework of clades. *International Journal of Plant Sciences*, 165: 7-21.
- Magoulick DA, Kobza RM. 2003. The role of refugia for fishes during drought: a review and synthesis. *Freshwater Biology*, 48: 1186-1198.
- Maia H, Maia C, Pires D, Alexandre V. 2006. Biology of the Iberian chub (*Squalius carolitertii*) in an atlantic-type

- stream (river Lima basin-north Portugal). A preliminary approach. *Limnética*, 25(3): 713-722.
- Mäkinen HS, Merilä J. 2008. Mitochondrial DNA Phylogeography of the three-spined stickleback (*Gasterosteus aculeatus*) in Europe – Evidence for multiple glacial refugia. *Molecular Phylogenetics and Evolution*, 46(1): 167-182.
- Mallet J, Besansky N, Hahn MW. 2015. How reticulated are species? *BioEssays*, 38(2): 140-149.
- Manel S, Schwartz MK, Luikart G, Taberlet P. 2003. Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology and Evolution*, 18: 189-197.
- Marková S, Šanda R, Crivelli A, Shumka S, Wilson IF, Vuković J, Berrebi P, Kotlík P. 2010. Nuclear and mitochondrial DNA sequence data reveal the evolutionary history of *Barbus* (Cyprinidae) in the ancient lake systems of Balkans. *Molecular Phylogenetics and Evolution*, 55(2): 488-500.
- Martin AP, Palumbi SR. 1993. Body size, metabolic rate, generation time, and the molecular clock. *Proceedings of the National Academy of Sciences of the United States of America*, 90(9): 4087-4091.
- Martínez-Capel F., García de Jalón D, Werenitzky D, Baeza D, Rodilla-Amallá M. 2009. Microhabitat use by three endemic Iberian cyprinids in Mediterranean rivers (Tagus River Basin, Spain). *Fisheries Management and Ecology*, 16(1): 52-60.
- Marton E. 2008. Paleomagnetism and palaeogeography. En: McCann T (ed.) *The geology of central Europe. Volumen 2: Mesozoic and Cenozoic*. Geological Society. London. Pp: 1022-1035.
- Mayden RL. 1988. Vicariance biogeography, parsimony, and evolution in North American freshwater fishes. *Systematic Biology*, 37(4): 329-355.
- Mayden RL. 1997. A hierarchy of species concept: The denouement in the saga of the species problem. En: Claridge MF, Dawah HA, Wilson MR (eds.). *Species: The units of Biodiversity*. Chapman and Hall. London., pp: 381-424.
- Mayr E. 1963. *Animal species and evolution*. Mass: Belknap Press. Cambridge. Harvad.
- Mayr E. 1982. *Speciation and Macroevolution*. *Evolution*, 36: 1119-1132.
- Médail F, Quézel P. 1999. Biodiversity hotspots in the Mediterranean Basin: setting global conservation priorities. *Conservation Biology*, 13(6): 1510-1513.
- Mendes J, Harris DJ, Carranza S, Salvi D. 2016. Evaluating the phylogenetic signal limit from mitogenomes, slow evolving genes, and the concatenation approach. New insights into the Lacertini radiation using fast evolving nuclear genes and species trees. *Molecular Phylogenetics and Evolution*, 100: 254-267.
- Mesquita N, Cunha C, Carvalho GR, Coelho MM. 2007. Comparative Phylogeography of endemic cyprinids in the southwest Iberian Peninsula: evidence for a new ichthyological area. *Journal of Fish Biology*, 71(sa): 45-75.
- Mesquita N, Hänfling B, Carvalho GR, Coelho MM. 2005. Phylogeography of the cyprinid *Squalius aradensis* and implications for conservation of the endemic freshwater fauna of southern Portugal. *Molecular Ecology*, 14(7): 1939-1954.
- Meulenkamp JE, Sissingh W. 2000. Map 17-23. Tertiary. En: Decourt J, Gaetani (eds.). *Atlas Peri-Tethys, Palaeogeographical maps*. CCGM/CGWM. Paris. pp: 153-268.
- Meybeck M. 1995. Global distribution of lakes. En: Lerman A, Imboden D, Gat J. *Physics and Chemistry of Lakes*. Springer, Berlin. pp: 1-32.
- Miller J. 2010. Species distribution modeling. *Geography Compass*, 4(6): 490-509.
- Miller KG, Kominz MA, Browning JV, Wright JD, Mountain GS, Katz ME, Sugarman PJ, Cramer BS, Christie-Blick N, Pekar SF. 2005. The Phanerozoic record of global sea-level change. *Science*, 310(5752): 1293-1298.
- Miresan V, Latiu C, Cocan D. 2016. Age structure and growth dynamics of European chub, *Squalius cephalus*, Linnaeus 1758 (Actinopterygii: Cyprinidae) from two catchments in Transylvania-Romania. *Bulletin UASVM Animal Science and Biotechnologies*, 73(2): 206-211.
- Mittelbach GG, Schemske DW, Cornell HV, Allen AP, Brown JM, Bush MK *et al.* 2007. Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. *Ecology Letters*, 10(4): 315-331.

- Moore WS. 1995. Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees. *Evolution*, 49(4): 718-726.
- Moritz C y Bermingham E. 1998. Comparative phylogeography: concepts and applications. *Molecular Ecology*, 7: 367-370.
- Muñoz JA, Martínez A, Vergés J. 1986. Thrust sequences in the eastern Spanish Pyrenees. *J. Struct. Geol.*, 8: 399-405.
- Myers GS. 1938. Fresh-water fishes and West Indian Zoogeography. *Annual Report of the Board of Regents of the Smithsonian Institution*, 92: 339-364.
- Myers N, Mittermeier RA, Mittermeier CG, de Fonseca GAB, Kent J. 2000. Biodiversity hotspots for conservation priorities. *Nature*, 403: 853-858.
- Naomi S-I. 2011. On the integrated frameworks of species concepts: Mayden's hierarchy of species concepts and de Queiroz's unified concept of species. *Journal of Zoological Systematics and Evolutionary Research*, 49(3): 177-184.
- Near TJ, Bernard MF. 2004. Rapid allopatric speciation in logperch darters (Percidae: *Percina*). *Evolution*, 58: 2798-2808.
- Near TJ, Keck BP. 2005. Dispersal, vicariance, and timing of diversification in *Nothotus* darters. *Molecular Ecology*, 14(11): 3485-3496.
- Near TJ, Meylan PA, Shaffer HB. 2005. Assessing concordance of fossil calibration points in molecular clock studies: an example using turtles. *American Naturalist*, 165: 137-146.
- Nei M, Maruyama T, Chakrabarty R. 1975. The bottleneck effect and genetic variability in populations. *Evolution*, 29(1): 1-10.
- Nelson JS, Grande TC y Wilson MVH. 2016. Fishes of the world. 5th edition. John Wiley & Sons, Hoboken, New Jersey.
- Nesbø CL, Fossheim T, Vøllestad LA. 1999. Genetic divergence and phylogeographic relationships among European perch (*Perca fluviatilis*) populations reflect glacial refugia and postglacial colonization. *Molecular Ecology*, 8(9): 1387-1404.
- Nikolsky GV. 1954. *Special ichthyology*. 2nd edition. Sovetskaya Nauka. Moscow. Russia.
- Nylander JAA, Wilgenbusch JC, Warren DL, Swofford DL. 2008. AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. *Bioinformatics*, 24(4): 581-583.
- Nix HA. 1986. A biogeographic analysis of Australian Elapid snakes. En: Longmore (ed.). *Australian flora and fauna. Series 7: Atlas of Elapid snakes of Australia*. Australian Government Publishing Service. Canberra. pp: 4-15.
- Norton W, Bally-Cuif L. 2010. Adult zebrafish as a model organism for behavioural genetics. *BMC Neuroscience*, 11:90.
- Ohno A, Muramoto J, Christian L, Atkin NB. 1967. Diploid-Tetraploid relationships among old-world members of the fish family Cyprinidae. *Chromosoma (Berl.)*, 23:1-9.
- Oikonomou A, Leprieur F, Leonardos ID. 2014. Biogeography of freshwater fishes of the Balkan Peninsula. *Hydrobiologia*, 738: 205-220.
- O'Meara BC. 2010. New heuristic methods for joint species delimitation and species tree inference. *Systematic Biology*, 59: 59-73.
- Paillard D. 1998. The timing of Pleistocene glaciations from a simple multiple-state climate model. *Nature*, 391: 378-381.
- Pamilo P, Nei M. 1988. Relationships between gene trees and species trees. *Molecular Biology and Evolution*, 5: 568-583.
- Papanikolaou D, Fountoulis I, Metaxas C. 2007. Active faults, deformation rates and Quaternary paleogeography at Kyprissiakos Gulf (SW Greece) deduced from onshore and offshore data. *Quaternary International*, 172: 14-30.

- Paramonova NP, Shcherbba IG, Khondikarian SO. 2004. MAP 7: Late Middle Miocene (Late Serravallian, Sarmatian s.s., Middle Sarmatian s.l.). En: Popov SV, Rögl F, Rozanov AY, Steininger FF, Scherba IG, Kovac M (eds.). *Lithological-Palaeogeographic maps of Paratethys*. Cour. Forsch. Inst. Senckenberg, 250: 27-29.
- Pavelic D, Belak M. 2008. Dinarids in north Croatia and Bosnia. En: McCann T (ed.). *The geology of central Europe. Volume 2: Mesozoic and Cenozoic*. Geological Society. London, pp: 1102-1116.
- Perdices A, Bermingham E, Montilla A, Doadrio I. 2002. Evolutionary history of the genus *Rhamdia* (Teleostei: Pimelodidae) in Central America. *Molecular Phylogenetics and Evolution*, 25: 172-189.
- Perdices A, Böhlen J, Doadrio I. 2008. The molecular diversity of Adriatic spined loaches (Teleostei: Cyprinidae). *Molecular Phylogenetics and Evolution*, 46: 382-390.
- Perdices A, Böhlen J, Šlechtová, Doadrio I. 2016. Molecular evidence for multiple origins of the European spined loaches (Teleostei, Cobitidae). *PLoS One*, 11(3): e0151228.
- Perdices A, Doadrio I, Economidisi PS, Böhlen J, Blunrescu P. 2003. Pleistocene effects on the European freshwater fish fauna: double origin of the cobitid genus *Sabanejewia* in the Danube basin (Osteichthyes: Cobitidae). *Molecular Phylogenetics and Evolution*, 26(2): 289-299.
- Perdices A, Doadrio I, Côté I, Machordom A. 2000. Genetic divergence and origin of Mediterranean populations of the river blenny *Salaria fluviatilis* (Teleostei: Blenniidae). *Copeia*, 2000(3): 723-731.
- Pereira LHG, Hanner R, Foresti F, Oliveira C. 2013. Can DNA barcoding accurately discriminate megadiverse Neotropical freshwater fish fauna?. *BMC Genetics*, 14:20.
- Pereira LHG, Pazian MF, Hanner R, Foresti F, Oliveira C. 2011. DNA barcoding reveals hidden diversity in the Neotropical freshwater fish *Piabina argentea* (Characiformes: Characidae) from the Upper Paraná Basin of Brazil. *Mitochondrial DNA*, 22(1): 87-96.
- Pereira SL, Baker AL, Wojtala A. 2002. Combined nuclear and mitochondrial DNA sequences resolve generic relationships within the Cracidae (Galliformes, Aves). *Systematic Biology*, 51: 946-958.
- Pérez-Rodríguez R, Domínguez-Domínguez O, Mar-Silva FA, Doadrio I, Pérez-Ponce de León G. 2016. The historical biogeography of the southern group of the sucker genus *Moxostoma* (Teleostei: Catostomidae) and the colonization of central Mexico. *Zoological Journal of the Linnean Society*, 177(3): 633-647.
- Phillips SJ et al., 2004. *A maximum entropy approach to species distribution modeling*. En: Proceedings of the Twenty-First International Conference on Machine Learning. ACM Press. New York., pp: 472-486.
- Phillips SJ et al., 2006. Maximum entropy modeling of species geographic distribution. *Ecol. Model.*, 190: 231-259.
- Phillips SJ, Dudík M. 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. 2008. *Ecography*, 31: 161-175.
- Phillips SJ, Dudík M, Elith J, Graham C, Lehman A, et al., 2009. Sample selection bias and presence-only models of species distributions. *Ecol. Appl.*, 19: 181-197.
- Picq S, Alda F, Krahe R, Bermingham E. 2014. Miocene and Pliocene colonization of the Central American isthmus by the weakly electric fish *Brachyhypopomus occidentalis* (Hypopomidae, Gymnotiformes). *Journal of Biogeography*, 41(8): 1520-1532.
- Poczai P, Hyvönen J. 2010. Nuclear ribosomal spacer regions in plant phylogenetics: problems and prospects. *Molecular Biology Reports*, 37(4): 1897-1912.
- Pollard DA, Iyer VN, Moses AM, Elsen MB. 2006. Widespread discordance of gene trees with species tree in *Drosophila*: evidence for incomplete lineage sorting. *PLoS Genetics*, 2(20): e173.
- Pons J, Barraclough TG, Gómez-Zurita J, Cardoso A, Durand DP, Hazell S, Kamoun S, Sumlin WD, Vogler AP. 2006. Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Systematic Biology*, 55: 595-609.
- Popov SV, Rögl F, Rozanov AY, Steininger FF, Shcherba IG, Kovac M (eds.). 2004. Lithological-Paleogeographical maps of Paratethys 10 maps Late Eocene to Pliocene. Courier Forschungsinstitut Senckenberg. 46 pp.

- Popov SV, Shcherba IG, Ilyna LB, Nevesskaya LA, Paramonova PA, Khondkarian SO, Magyar I. 2006. Late Miocene to Pliocene paleogeography of the Paratethys and its relation to the Mediterranean. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 238(1-4): 91-106.
- Posada D, Crandall KC. 2001. Intraspecific gene genealogies: trees grafting into networks. *Trends in Ecology and Evolution*, 16: 37-45.
- Provan J, Bennett KD. 2008. Phylogeographic insights into cryptic glacial refugia. *Trends in Ecology and Evolution*, 23(10): 564-571.
- Puigdefábregas C, Muñoz JA, Vergés J. 1992. En: McClay KR (ed.). Thrusting and foreland basin evolution in the southern Pyrenees. *Thrust tectonics*. Chapman & Hall. London, pp: 247-254.
- Pyron M. 1996. Sexual size dimorphism and phylogeny in North American minnows. *Biological Journal of the Linnean Society*, 57: 327-341.
- Rabosky DL y McCune AR. 2010. Reinventing species selection with molecular phylogenies. *Trends in Ecology and Evolution*, 25(2): 68-74.
- Rambaut A, Suchard M, Xie W, Drummond AJ. 2014. Tracer v.1.6. Institute of Evolutionary Biology, University of Edinburgh (disponible en <http://beast.bio.ed.ac.uk/>).
- Ratschbacher L, Merle O, Davy P y Cobbold P. 1991. Lateral extrusion in the eastern Alps. Part I: boundary conditions and experiments scaled for gravity. *Tectonics*, 10: 245-256.
- Ree RH, Smith SA. 2008. Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology*, 57: 4-14.
- Reisz RR, Müller J. 2004. Molecular timescales and the fossil record: a paleontological perspective. *Trends in Genetics*, 20: 237-241.
- Reyjol Y, Hugueny B, Pont D, Bianco PG, Beier U, Caiola N, Casals F, Cowx I, Economou A, Ferreira T, Haidvogel G, Noble R, De Sostoa A, Vigneron T, Virbickas T. 2008. Patterns in species richness and endemism of European freshwater fish. *Global Ecology and Biogeography*, 16(1): 65-75.
- Ribeiro F, Cowx IG, Tiago P, Filipe AF, Moreira da Costa L, Collares-Pereira MJ. 2003. Growth and reproductive traits of diploid and triploid forms of the *Squalius alburnoides* cyprinid complex in a tributary of the Guadiana River, Portugal. *Archiv für Hydrobiologie*, 156(4): 471-484.
- Ullian O, Zardoya R, Doadrio I. 2008. Phylogenetic relationships of Middle American cichlids (Cichlidae, Heroini) based on combined evidence from nuclear genes, mtDNA, and morphology. *Molecular Phylogenetics and Evolution*, 49(3): 941-957.
- Rissler LJ, Apodaca JJ, Wiens J. 2007. Adding more ecology into species delimitation: ecological niche models and phylogeography help define cryptic species in the black salamander (*Aneides flavipunctatus*). *Systematic Biology*, 56(6): 924-942.
- Robba L, Russell SJ, Barker GL y Brodie J. 2006. Assessing the use of the mitochondrial cox1 marker for use in DNA barcoding of red algae (Rhodophyta). *American Journal of Botany*, 93(8): 1101-1108.
- Robalo JI, Almada VC, Levy A, Doadrio I. 2007a. Re-examination and phylogeny of the genus *Chondrostoma* based on mitochondrial and nuclear data and the definition of 5 new genera. *Molecular Phylogenetics and Evolution*, 42(2): 362-372.
- Robalo JI, Sousa-Santos C, Levy A, Almada VC. 2007b. Molecular insights on the taxonomic position of the paternal ancestor of the *Squalius alburnoides* hybridogenetic complex. *Molecular Phylogenetics and Evolution*, 39(1): 276-281.
- Rokas A, Williams BL, King N, Carroll SB. 2003. Genome-scale approaches to resolving incongruence in molecular phylogenies. *Nature*, 425(6960): 798-804.
- Rögl F. 1999. Circum-Mediterranean Miocene Palaeogeography. En: Rössner GE, Heissig K (eds.). *The Miocene land mammals of Europe*. Verlag, Pr Friedrich. Pfeil. München: 9-24.

- Rögl F. 2001. Circum-Mediterranean Miocene Palaeogeography. *Ver. Inst. Geol. Paleontol. Karl-Franzens Univ Graz Oesterr*, 4: 49-59.
- Ronquist F. 1997. Dispersal-Vicariance analysis: a new approach to the quantification of historical biogeography. *Systematic Biology*, 46(1): 195-203.
- Rüber L, Kottelat M, Tan HH, Ng PK y Britz R. 2007. Evolution of miniaturization and the phylogenetic position of *Paedocypris* comprising the world's smallest vertebrate. *BMC Evolutionary Biology*, 7: 38.
- Rutschmann F. 2006. Molecular dating of phylogenetic trees: a brief review of current methods that estimate divergence times. *Diversity and Distributions*, 12: 35-48.
- Rutschmann F, Eriksson T, Abu Salim K, Conti E. 2007. Assessing calibration uncertainty in molecular dating: the assignment of fossils to alternative calibration points. *Systematic Biology*, 56: 591-608.
- Ruzzante DE, Walde SJ, Gosse JC, Cussac VE, Habit E, Zemlak TS, Adams EDM. 2008. Climate control on ancestral populations dynamics: insight from Patagonian fish phylogeography. *Molecular Ecology*, 17(9): 2234-2244.
- Saitoh K, Sado T, Mayden RL, Hanzawa N, Nakamura K, Nishida M y Miya M. 2006. Mitogenomic evolution and interrelationships of the Cypriniformes (Actinopterygii, Ostariophysi): the first evidence towards resolution of higher-level relationships of the world's largest fish family clade based on 59 whole mitogenome sequences. *Journal of Molecular Evolution*, 63: 826-841.
- Saitoh et al., 2011. Evidence from mitochondrial genomics supports the Lower Mesozoic of South Asia as the time and place of basal divergence of cypriniform fishes (Actinopterygii: Ostariophysi). *Zoological Journal of the Linnean Society*, 19: 1291-1303.
- Sakai H, Goto A, Jeon S-R. 2002. Speciation and dispersal of *Tribolodon* species (Pisces, Cyprinidae) around the Sea of Japan. *Zoological Science*, 19(11): 1291-1303.
- Sakai H, Ito Y, Shedko SN, Frolov SV, Chereshev IA, Jeon S-R, Goto I. 2006. Phylogenetic and taxonomic relationships of northern far eastern phoxin minnows, *Phoxinus* and *Rhynchocypris* (Pisces, Cyprinidae), as inferred from allozyme and mitochondrial 16S rRNA sequence analyses. *Zoological Science*, 23(4): 323-331.
- Salichos L, Rokas A. 2013. Inferring ancient divergences requires genes with strong phylogenetic signal. *Nature*, 497(7449): 327-331.
- Salzburger W, Brandstätter A, Gilles A, Parson W, Hempel M, Sturmbauer C, Meyer A. 2003. Phylogeography of the vairone (*Leuciscus souffia*, Risso, 1826) in central Europe. *Molecular Ecology*, 12(9): 2371-2386.
- Sánchez-Hernández J, Cobo F. 2011. Summer food resource partitioning between four sympatric fish species in Central Spain (River Tormes). *Folia Zoologica*, 60(3): 189-202.
- Šanda R, Vukić J, Choleva L, Krozek J, Šediva A, Shumka S, et al., 2008. Distribution of loach fishes (Cobitidae, Nemachilidae) in Albania, with genetic analysis of populations of *Cobitis ohridana*. *Folia Zoologica*, 57(1-2): 42-50.
- Sanderson MJ. 1997. A nonparametric approach to estimating divergence times in the absence of rate constancy. *Molecular Biology and Evolution*, 14: 1218-1231.
- Sanderson MJ. 2002. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. *Molecular Biology and Evolution*, 19: 101-109.
- Sanger F, Coulson AR. 1975. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *Journal of Molecular Evolution*, 94: 441-448.
- Sanger F, Nicklen S, Coulson AR. 1977. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America*, 94: 441-448.
- Sanjurjo OI, Carmona JA, Doadrio I. 2003. Evolutionary and biogeographical patterns within Iberian populations of the genus *Squalius* inferred from molecular data. *Molecular Phylogenetics and Evolution*, 29(1): 20-30.
- Sanmartín I. 2001. Patterns of animal dispersal, vicariance and diversification in the Holarctic. *Biological Journal of*

- the *Linnean Society*, 73(4): 345-390.
- Sanmartín I. 2003. Dispersal vs vicariance in the Mediterranean: historical biogeography of the Palearctic *Pachydominae* (Coleoptera, Scarabaeoidea). *Journal of Biogeography*, 30: 1883-1897.
- Sasaki T, Kartavtsev YP, Chiba SN, Uematsu T, Sviridov VV y Hanzawa N. 2007. Genetic divergence and phylogenetic independence of Far Eastern species in subfamily Leuciscinae (Pisces: Cyprinidae) inferred from mitochondrial DNA analyses. *Genes and Genetic Systems*, 82: 329-340.
- Sathiamurthy E, Voris HK. 2006. Maps of Holocene Sea Level Transgression and submerged lakes on the Sunda shelf. *The Natural History Journal of Chulalongkorn University*, 2: 1-44.
- Schlick-Steiner BC, Steiner FM, Seifert B, Stauffer C, Christian E y Crozier RH. 2010. Integrative taxonomy: a multisource approach to exploring biodiversity. *Annual Review of Entomology*, 55: 421-438.
- Schmitt T. 2007. Molecular biogeography of Europe: Pleistocene cycles and postglacial trends. *Frontiers in Zoology*, 4:11.
- Schultz RJ. 1980. The role of polyploidy in the evolution of fishes. En: Lewis WH (ed.). *Polyploidy: biological relevance*. New York. Plenum Press, pp: 319-339.
- Schuster SC. 2008. Next-generation sequencing transforms today's biology. *Nature Methods*, 5(1): 16-18.
- Seager R, Kiu H, Henderson N, Simpson I, Kelley C, Shaw T, Kushnir Y, Ting M. 2014. Causes of increasing aridification of the Mediterranean region in response to rising greenhouse gases. *Journal of Climate*, 27: 4655-4676.
- Seehausen O. 2004. Hybridization and adaptive radiation. *Trends in Ecology and Evolution*, 19849: 198-207.
- Seifertová M, Bryja J, Vyskočilová M, Martínková N, Šimková A. 2012. Multiple Pleistocene refugia and post-glacial colonization in the European chub (*Squalius cephalus*) revealed by combined use of nuclear and mitochondrial markers. *Journal of Biogeography*, 39(6): 1024-1040.
- Shapiro B, Ho SYW. 2004. Ancient hyaenas highlight the old problem of estimating evolutionary rates. *Molecular Ecology*, 23: 499-501.
- Shubin N, Tabin C y Carroll S. 2009. Deep homology and the origins of evolutionary novelty. *Nature*, 457: 818-823.
- Singh G, Dal Grande F, Divakar PK, Otte J, Leavitt SD, Szczepanska K, Crespo A, Rico VJ, Aptroot A, Cáceres MES, Lumbsch HT, Schmitt I. 2015. Coalescent-based species delimitation approach uncovers high cryptic diversity in the Cosmopolitan lichen-forming fungal genus *Protoparmelia* (Lecanorales, Ascomycota). *PLoS One*, 10(5): e0124625.
- Siroco F, Seelos K, Schaber K, Rein B, Deher F, Diehl M, Lehne R, Jagër K, Krbertschek M, Degering D. 2005. A late Eemian aridity pulse in central Europe during the last glacial inception. *Nature*, 436: 833-836.
- Simons AM, Mayden RL. 1998. Phylogenetic relationships of the Western North American Phoxinini (Actinopterygii: Cyprinidae) as inferred from mitochondrial 12S and 16S ribosomal RNA sequences. *Molecular Phylogenetics and Evolution*, 9(2): 308-329.
- Šlechtová V, Böhlen J, Freyhof J, Persat H, Delmastro GB. 2004. The Alps as barrier to dispersal in cold-adapted freshwater fishes? Phylogeographic history and taxonomic status of the bullhead in the Adriatic freshwater drainage. *Molecular Phylogenetics and Evolution*, 33(1): 225-239.
- Šlechtová V, Böhlen J, Perdices A. 2007. Molecular phylogeny of the freshwater fish family Cobitidae (Cypriniformes: Teleostei): Delimitation of genera, mitochondrial introgression and evolution of sexual dimorphism. *Molecular Phylogenetics and Evolution*, 47: 812-831.
- Smith KG, Darwall WR (eds.). 2006. The status and distribution of freshwater fish endemic to the Mediterranean Basin (Vol. 1). IUCN.
- Song S, Liu L, Edwards SV, Wu S. 2012. Resolving conflict in eutherian mammal phylogeny using phylogenomics and the multiple coalescent model. *Proceedings of the National Academy of Sciences of the United States of America*, 109(37): 14942-14947.

- Sota T, Vogler AP. 2001. Incongruence of mitochondrial and nuclear gene trees in the Carabid beetle *Ohomopterus*. *Systematic Biology*, 50(1): 39-59.
- Sousa-Santos C, Collares-Pereira MJ, Almada VC. 2006. Evidence of extensive mitochondrial introgression with nearly complete substitution of the typical *Squalius pyrenaicus*-like mtDNA of the *Squalius alburnoides* complex (Cyprinidae) in an independent Iberian drainage. *Journal of Fish Biology*, 68(SB): 292-301.
- Sousa-Santos C, Collares-Pereira MJ, Almada VC. 2007. Reading the history of a hybrid fish complex from its molecular record. *Molecular Phylogenetics and Evolution*, 45(3): 981-996.
- Spielman D, Brook BB, Frankham R. 2004. Most species are not driven to extinction before genetic factors impact them. *Proceedings of the National Academy of Sciences*, 101: 15261-15264.
- Spink PQ, Shaffer HB. 2009. Conflicting mitochondrial and nuclear phylogenies for the widely disjunct *Emys* (Testudines: Emydidae) species complex, and what they tell us about biogeography and hybridization. *Systematic Biology*, 58(1): 1-20.
- Sprovieri M, Sacchi M, Rohling EJ. 2003. Climatically influenced interaction between the Mediterranean and the Paratethys during Tortonian. *Paleoceanography*, 18(2): 1034.
- Stankovic S. 1960. The Balkan lake Ohrid and its living world. En: Bodenheimer FS, Weisbach WW (eds.). *Monographiae Biologicae IX*. Den Haag, The Netherlands.
- Stefani F, Galli P, Crosa G, Zaccara S, Calamari D. 2004. Alpine and Apennine barriers determining the differentiation of the rudd (*Scardinius erythrophthalmus* L.) in the Italian peninsula. *Ecology of Freshwaters*, 13: 168-175.
- Stockwell DRB, Peters DP. 1999. The GARP modelling system: Problems and solutions to automated spatial prediction. *International Journal of Geographical Information Science*, 13(2): 143-158.
- Storfer A, Murphy MA, Evans JS, Goldberg CS, Robinson S, Spear SF, Dezzani R, Delmelle E, Vierling L, Waits LP. 2007. Putting the "landscape" in landscape genetics. *Heredity*, 98: 128-142.
- Storfer A, Murphy MA, Spear SF, Waits LP. 2010. Landscape genetics: where are we now?. *Molecular Ecology*, 19(17): 3496-3514.
- Stoumboudi MTh, Kottelat M, Barbieri R. 2006. The fishes of the inland waters of Lesbos Island, Greece. *Ichthyological Exploration of Freshwaters*, 17(2): 129-146.
- Strange RM y Mayden RL. 2009. Phylogenetic relationships and a revised taxonomy for North American cyprinids currently assigned to *Phoxinus* (Actinopterygii: Cyprinidae). *Copeia*, 2009: 494-501.
- Suc J-P. 1984. Origin and evolution of the Mediterranean vegetation and climate in Europe. *Nature*, 307: 429-432.
- Suc J-P, Zagwijn WH. *Boreas*, 12: 153-166.
- Šušnik S, Snoj A, Dovc P. 2001. Evolutionary distinctness of grayling (*Thymallus thymallus*) inhabiting the Adriatic river system, as based on mitochondrial DNA variation. *Biological Journal of the Linnean Society*, 74(3): 375-385.
- Swartz ER, Chakona A, Skelton PH, Bloomer P. 2007. The genetic legacy of lower sea levels: does the confluence of rivers during the last glacial maximum explain the contemporary distribution of a primary freshwater fish (*Pseudobarbus burchellii*, Cyprinidae) across isolated river systems? *Hydrobiologia*, 726(1): 109-121.
- Taanam J-W. 1999. The mitochondrial genome: structure, transcription, translation and replication. *Biochemica et Biophysica Acta (BBA) – Bioenergetics*. 1410(2): 103-123.
- Tang Q-Y, Liu S-Q, Yu D, Liu H-Z, Danley PD. 2012. Mitochondrial capture and incomplete lineage sorting in the diversification of balitorine loaches (Cypriniformes, Balitoridae) revealed by mitochondrial and nuclear genes. *Zoologica Scripta*, 41(3): 233-247.
- Tang Q, Xiong B, Lium H. 2005. Phylogeny of the East Asian botiine loaches (Cypriniformes, Botiidae) inferred from mitochondrial cytochrome *b* gene sequences. *Hydrobiologia*, 544: 249-258.
- Tautz D, Arctander P, Minelli A, Thomas RH y Vogler AP. 2003. A plea for DNA taxonomy. *Trends in Ecology and Evolution*, 18(2): 70-74.

- Talavera G, Vila R. 2011. What is the phylogenetic signal limit from mitogenomes? The reconciliation between mitochondrial and nuclear data in the Insecta class phylogeny. *BMC Evolutionary Biology*, 11: 315.
- Templeton AR. 2006. *Population genetics and microevolutionary theory*. John Wiley & Sons. New Jersey, USA. 720 pp.
- Thorne JL, Kishino H, Painter IS. 1998. Estimating the rates of evolution of the rate of molecular evolution. *Molecular Biology and Evolution*, 15: 1647-1657.
- Ting N, Tosi AJ, Li Y, Zhang Y-P, Disotell TR. 2008. Phylogenetic incongruence between nuclear and mitochondrial markers in the Asian colobines and the evolution of the langurs and leaf monkeys. *Molecular Phylogenetics and Evolution*, 46(2): 466-474
- Travers RA. 1989. Systematic account of collection of fishes from the Mongolian People's Republic: with a review of the hydrobiology of the major Mongolian drainage basins. *Bull. Br. Mus. Nat. Hist. Zool.*, 55: 173-207.
- Tsigenopoulos CS, Durand J-D, Unlü E, Berrebi P. 2003. Rapid radiation of the Mediterranean *Barbus* species (Cyprinidae) after the Messinian salinity crisis of the Mediterranean Sea, inferred from mitochondrial phylogenetic analyses. *Zoological Journal of the Linnean Society*, 80(2): 207-222.
- Tsigenopoulos CS, Karakousis Y, Berrebi P. 1999. The North Mediterranean *Barbus* lineage: phylogenetic hypotheses and taxonomic implications based on allozymic data. *Journal of Fish Biology*, 54(2): 267-286.
- Turan D, Kottelat M, Bayçelebi E. 2017. *Squalius semae*, a new species of chub from the Euphrates River, Eastern Anatolia (Teleostei: Cyprinidae). *Zoology in the Middle East*, atolia (Teleostei: Cyprinidae), *Zoology in the Middle East*, DOI: 10.1080/09397140.2017.1290761.
- Unmack PJ, Hämmner MP, Adams M, Dowling TE. 2001. A phylogenetic analysis of pygmy perches (Teleostei: Percichthyidae) with an assessment of the major historical influences on aquatic biogeography in southern Australia. *Systematic Biology*, 60(6): 797-812.
- Uyeno T, Smith GR. 1972. Tetraploid origin of the karyotype of catostomid fishes. *Science*, 175: 644-646.
- Van Couvering JAH. 1977. Early records of freshwater fishes in Africa. *Copeia*, 163-166.
- Van Dam JA y Weltje GJ. 1999. Reconstruction of the Late Miocene climate of Spain using rodent paleocommunity successions: an application of end-member modeling. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 151: 267-305.
- Vences M, Thomas M, Bonnet RM y Vieites DR. 2005. Deciphering amphibian diversity through DNA barcoding: chances and challenges. *Philosophical Transactions of the Royal Society B*. 360: 1859-1868.
- Wagner B, Vogel H, Zanchetta G, Sulpizio R. 2010. Environmental change within the Balkan region during the past ca. 50 ka recorder in the sediments from lakes Prespa and Ohrid. *Biogeosciences*, 7: 3187-3198.
- Wang XZ, Gan XN, Li JB, Mayden RL y He SP. 2012. Cyprinid phylogeny based on Bayesian and maximum likelihood analyses of partitioned data: implications for Cyprinidae systematics. *Science China Life Sciences*, 55(9): 761-733.
- Wang X, Li J y He S. 2007. Molecular evidence for the monophyly of East Asian groups of Cyprinidae (Teleostei: Cypriniformes) derived from the nuclear recombination activating gene 2 sequences. *Molecular Phylogenetics and Evolution*, 42: 157-170.
- Ward RD, Costa FO, Holmes BH y Steinke D. 2008. DNA barcoding of shared fish species from the North Atlantic and Australasia: minimal divergence for most taxa, but *Zeus faber* and *Lepidopus caudatus* each probably constitute two species. *Inter-Research Aquatic Biology*, 3: 71-87.
- Ward RD y Holmes BH. 2007. An analysis of nucleotide and amino acid variability in the barcode region of cytochrome c oxidase I (*cox1*) in fishes. *Molecular Ecology Notes*, 7: 899-907.
- Waters JM, Craw D, Yungson JH, Wallis GP. 2001. Genes meet geology: fish phylogeographic pattern reflects ancient, rather than modern, drainage connections. *Evolution*, 55: 1844-1851.
- Waters JM, Dijkstra LH, Wallis G. 2000. Biogeography of a southern hemisphere freshwater fish: how important is

- marine dispersal?. *Molecular Ecology*, 9: 1815-1821.
- Whittle CA, Johnston MO. 2003. Broad-scale analysis contradicts the theory that generation time affects molecular evolutionary rates in plants. *Journal of Molecular Evolution*, 56: 223-233.
- Wiens JJ. 2004. Speciation and ecology revisited: phylogenetic niche conservatism and the origin of species. *Evolution*, 58: 193-197.
- Wiens JJ y Donoghue MJ. 2004. Historical biogeography, ecology and species richness. *Trends in Ecology and Evolution*, 19: 639-644.
- Wiens JJ, Kuczynski CA, Smith SA, Mulcaty DG, Sites JW Jr, Townsend TM, Reeder TW. 2008. Branch lengths support and congruence: testing the phylogenomic approach with 20 nuclear loci I snakes. *Systematic Biology*, 57: 420-431.
- Wiley EO. 1978. The evolutionary species concept reconsidered. *Systematic Biology*, 27: 17-26.
- Wiley EO, Mayden RL. 2000. The evolutionary species concept. En: Wheeler QD, Meiner R (eds.). *Species Concept and Phylogenetic Theory: A debate*. Columbia University Press. New York., pp: 70-89.
- Winfield IJ y Nelson JE (eds.). 1991. *Cyprinid fishes: Systematics, Biology and Exploitation*. Chapman and Hall. London.
- Wysor B, Kooistra WHCF, Frederiq S. 2012. Comparative phylogeography of reticulate Cladophoralean algae. *Journal of Phycology*, 38(s1): 38-39.
- Wolf U, Ritter H, Atkin NB, Ohno S. 1969. Polyploidization in the fish family Cyprinidae, order Cypriniformes. *Humangenetik*, 7: 240-244.
- Woodburne MO. 2010. The Great American Biotic Interchange: dispersals, tectonics, climate, sea level and holding pens. *J. Mammal. Evol.*, 17: 245-264.
- Xiao W, Zhang Y, Liu H. 2001. Molecular Systematics of Xenocyprininae (Teleostei, Cyprinidae): Taxonomy, Biogeography, and Coevolution of a special group restricted in East Asia. *Molecular Phylogenetics and Evolution*, 18(2): 163-173.
- Yang L, Sado T, Hirt MV, Pasco-Viel E, Arunachalam M, Li J, Wang X, Freyh J, Saitoh K, Simons AM, Miya M, He S, Mayden RL. 2015. Phylogeny and polyploidy: Resolving the classification of cyprinine fishes (Teleostei: Cypriniformes). *Molecular Phylogenetics and Evolution* 85: 97-116.
- Yu Y, Harris AJ, He XJ. 2010. S-DIVA (statistical dispersal-vicariance analysis): a tool for inferring biogeographic histories. *Molecular Phylogenetics and Evolution*, 56(2): 848-850.
- Yu Y, Harris AJ, Blair C, He XJ. 2015. RASP (Reconstruct Ancestral State in Phylogenies): A tool for historical biogeography. *Molecular Phylogenetics and Evolution*, 87: 46-49.
- Zachos J, Pagani M, Sloan L, Thomas E, Billups K. 2001. Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science*, 292: 686-693.
- Zardoya R y Doadrio I. 1998. Phylogenetic relationships of Iberian Cyprinidae: Systematic and biogeographical implications. *Proceedings of the Royal Society of London Series B.*, 265: 1365-1372.
- Zardoya R y Doadrio I. 1999. Molecular evidence on the evolutionary and biogeographical patterns of European cyprinids. *Journal of Molecular Evolution*, 13: 122-131.
- Zemlak TS, Habit EM, Walde SJ, Battini MA, Adams EDM, Ruzzante DE. 2008. Across the southern Andes on fin: glacial refugia, drainage reversals and a secondary contact zone revealed by the phylogeographical signal of *Galaxian petei* in Patagonia. *Molecular Ecology*, 17(23): 5049-5061.
- Zeng Y, Liu H. 2009. Morphological variation of pharyngeal bones and teeth in the subfamily Gobioninae (Cypriniformes: Cyprinidae) and its functional adaptations. *Zoological Research*, 30: 699-706.
- Zogaris S, Economou AN, Dimopoulos P. 2009. Ecoregions in the southern Balkans: should their boundaries be revised? *Environmental Management*, 43: 682-697.
- Zuckerkandl E, Pauling L. 1962. Molecular disease, evolution, and genetic heterogeneity. En: Kasha M, Pullman B

- (eds.). *Horizons in Biochemistry*. Academic Press, New York., pp: 189-225.
- Zuckerkandl E, Pauling L. 1965. Evolutionary divergence and convergence in proteins. En: Bryson V, Vogel J. *Evolving genes and Proteins*. Academic Press, New York., pp: 97-166.

II. OBJETIVOS

La presente tesis se ha dividido en varios capítulos en los cuales se exponen una serie de casos de estudio que recogen diferentes procesos evolutivos que han ocurrido las especies mediterráneas del género *Squalius* Bonaparte, 1837 (Actinopterygii, Cyprinidae) y los factores históricos o contemporáneos responsables de dichos procesos. La diversidad específica del género *Squalius* es muy amplia (52 especies formalmente descritas), al igual que el área de distribución que ocupa el género en su conjunto (región Mediterránea y área occidental de Oriente Próximo), por lo que estudiar todos los mecanismos implicados en su evolución es imposible dentro del marco de esta tesis, de manera que se han escogido unos pocos ejemplos generales representativos de los diferentes procesos evolutivos que se citan en el género. Además, se han elegido las especies mediterráneas porque constituyen un linaje evolutivo independiente y antiguo dentro de la región Mediterránea.

Los tres primeros capítulos de la tesis se centran en resolver distintas cuestiones evolutivas específicas dentro del género relacionadas con la evolución paleogeográfica de la región Mediterránea y su influencia en la historia evolutiva de las especies de *Squalius* consideradas. De esta forma, en los capítulos I, II y III se consideran los distintos factores históricos que han intervenido en los procesos de especiación y diversificación de distintos linajes genéticos de *Squalius*. Así, El **Capítulo I** analiza la influencia del tectonismo en la diversificación del género *Squalius* en el sur de la Península Ibérica, un área extremadamente compleja y activa desde el punto de vista geológico y tectónico. En el **Capítulo II** se completa la historia filogeográfica de la especie *S. pyrenaicus*, la que presenta mayor área de distribución en la Península Ibérica, junto con la de la especie *S. carolitertii*. Por último, el **Capítulo III** trata sobre la influencia de la geomorfología del Karst de los Alpes Dináricos en la reciente especiación y rápida diversificación que han sufrido las especies del género pertenecientes al linaje Mediterráneo que habitan la vertiente oriental del Mar Adriático.

Por su parte, el **Capítulo IV** introduce la variable climática en historia evolutiva de los peces de agua dulce, y por tanto se evalúa su influencia sobre las poblaciones de una especie concreta del género *Squalius*, *Squalius valentinus*; en este capítulo se abordan metodologías de modelaje de idoneidad de hábitat y se discute su relación con la diversidad y estructura genética de las poblaciones de esta especie.

En el **Capítulo V** se trata la problemática derivada de la hibridación entre especies del género *Squalius*, con un ejemplo concreto en el Peloponeso (Grecia). Este capítulo pone de manifiesto la presencia de discordancias mitonucleares y su influencia en la reconstrucción de un árbol de especies correcto.

En resumen, los objetivos parciales de la presente tesis son los siguientes:

- 1.- Análisis de varias especies ibéricas, *S. carolitertii*, *S. malacitanus*, *S. pyrenaicus* y *S. valentinus*, con el fin filogenético de identificar linajes evolutivos dentro de la península ibérica. Este objetivo parcial se trata en los capítulos I y II.
- 2.- Análisis filogeográfico de varias especies ibéricas, *S. carolitertii*, *S. malacitanus*, *S. pyrenaicus* y *S. valentinus*, con el fin de identificar patrones de estructura y diversidad genética en estas especies y los procesos generadores de dichos patrones. Este objetivo parcial se trata en los capítulos I, II y III.
- 3.-Reconstrucción de áreas ancestrales y datación mediante un reloj molecular de las especies ibéricas del género *Squalius* para obtener los patrones biogeográficos de varias especies de este género en una de las áreas más antigua y tectónicamente más compleja de la región Mediterránea. Este objetivo parcial se trata en los capítulos I y II
- 4.- Análisis filogenético y datación de varias especies del género *Squalius* de los Balcanes, *S. illyricus*, *S. microlepis*, *S. svallize*, *S. tenellus* y *S. zрманjae*, con el fin de evaluar la influencia del karst dinárico sobre los procesos de especiación de algunas especies balcánicas del género. Este objetivo parcial se trata en el capítulo III.
- 5- Análisis filogeográfico de varias especies del género *Squalius* de los Balcanes, *S. illyricus*, *S. microlepis* y *S. svallize*, con el fin el fin de evaluar la influencia del karst dinárico sobre los patrones de estructura genética hallados en estas especies. Este objetivo parcial se trata en el capítulo III.
- 6.- Análisis morfológico y filogenético para evaluar una discordancia mitonuclear debida a introgresión entre dos especies griegas del género *Squalius*. Este objetivo parcial se trata en el capítulo V.

Cenozoic and climatic events in southern Iberian Peninsula: implications for the evolutionary history of freshwater fish of the genus *Squalius* (Actinopterygii, Cyprinidae)

Perea S, Cobo-Simon M y Doadrio I. 2016. Cenozoic tectonic and climatic events in southern Iberian Peninsula: implications for the evolutionary history of freshwater fishes of the genus *Squalius* (Actinopterygii, Cyprinidae). *Molecular Phylogenetics and Evolution*, 97: 155-169.
DOI: <http://dx.doi.org/10.1016/j.ymepv.2016.01.07>

RESUMEN

Los ecosistemas acuáticos del sur de la Península Ibérica se ubican en el borde entre la placa europea y la africana y representan una región muy compleja desde el punto de vista tectónico que se extiende a lo largo de varias edades geológicas, desde el levantamiento de la cordillera Bética en los periodos Serravaliense y Tortoniense hasta el presente. Esta área también ha estado sometida a la influencia del cambio de las condiciones climáticas desde el Plioceno Medio-Superior, cuando los patrones climáticos estacionales típicos del clima mediterráneo se establecieron. Consecuentemente, la ictiofauna del sur de la Península Ibérica constituye un interesante modelo de estudio para analizar la influencia de los eventos tectónicos y climáticos en su historia evolutiva. Los ciprínidos *Squalius malacitanus* y *Squalius pyrenaicus* se distribuyen alopátricamente en el sur de la Península Ibérica y su historia evolutiva puede haber sido definida por eventos tectónicos y climáticos Cenozoicos. En este estudio analizamos los genes *MT-CYB* (510 individuos) y *RAG1* (140 individuos) de ambas especies para reconstruir las relaciones filogenéticas y estimar los tiempos de divergencia y las áreas de distribución ancestrales de las dos especies y de sus poblaciones. También evaluamos sus niveles de estructura genética así como la cantidad de flujo génico entre sus poblaciones. Para investigar factores recientes paleogeográficos y climáticos afectando al sur de la Península Ibérica se modelaron cambios en el nivel del mar a través del tiempo, desde el LGM hasta el presente. Los análisis filogenéticos y de estructura geográfica y poblacional revelaron la existencia de dos especies bien soportadas en el sur de la Península Ibérica (*S. malacitanus* y *S. pyrenaicus*) y dos subclados (Atlántico y Mediterráneo) dentro de la especie *S. malacitanus*. El origen de *S. malacitanus* y la separación de sus poblaciones atlánticas y mediterráneas tuvo lugar durante los periodos Serravaliense/Tortoniense y Mioceno/Plioceno respectivamente. Estos eventos de divergencia ocurrieron durante el Plioceno Medio y el Pleistoceno en *S. pyrenaicus*. En ambas especies las poblaciones de las cuencas atlánticas presentaron valores mayores de diversidad genética que los de las cuencas mediterráneas, hecho que puede ser explicado por la antigua laguna de la Janda. El aislamiento de *S. malacitanus* fue más temprano y relacionado con el levantamiento de la cordillera Bética. La divergencia de sus poblaciones atlánticas y mediterráneas estuvo asociada con la creación de los sistemas de agua dulce del sur de la Península Ibérica próximos al Estrecho de Gibraltar. La presencia de *S. pyrenaicus* en el sur ibérico puede ser el resultado de una colonización reciente asociada a una captura fluvial, como demuestra nuestra reconstrucción biogeográfica.

Palabras clave: Reconstrucción de áreas ancestrales, Cenozoico, Filogenia, Sur de Iberia, Reloj molecular, *Squalius*.

ABSTRACT

Southern Iberian freshwater ecosystems located at the border between the European and African plates represent a tectonically complex region spanning several geological ages, from the uplifting of the Betic Mountains in the Serravalian-Tortonian periods to the present. This area has also been subjected to the influence of changing climatic conditions since Middle-Upper Pliocene when seasonal weather patterns were established. Consequently, the ichthyofauna of southern Iberia is an interesting model system for analysing the influence of Cenozoic tectonic and climatic events on its evolutionary history. The cyprinids *Squalius malacitanus* and *Squalius pyrenaicus* are allopatrically distributed in southern Iberia and their evolutionary history may have been defined by Cenozoic tectonic and climatic events. We analysed *MT-CYB* (510 specimens) and *RAG1* (140 specimens) genes of both species to reconstruct phylogenetic relationships and to estimate divergence times and ancestral distribution ranges of the species and their populations. We also assessed their levels of genetic structure and diversity as well as the amount of gene flow between populations. To investigate recent palaeogeographical and climatic factors in southern Iberia, we modelled changes-through-time in sea level from the LGM to the present. Phylogenetic, geographic and population structure analyses revealed two well-supported species (*S. malacitanus* and *S. pyrenaicus*) in southern Iberia and two subclades (Atlantic and Mediterranean) within *S. malacitanus*. The origin of *S. malacitanus* and the separation of its Atlantic and Mediterranean populations occurred during the Serravalian-Tortonian and Miocene-Pliocene periods, respectively. These divergence events occurred in the Middle Pliocene and Pleistocene in *S. pyrenaicus*. In both species, Atlantic basins possessed populations with higher genetic diversity than Mediterranean, which may be explained by the Janda Lagoon. The isolation of *S. malacitanus* was earlier and related to the rising of the Betic Mountains. Divergence of its Atlantic and Mediterranean populations was associated with the creation of the freshwater systems of southern Iberia close to the Gibraltar Strait. The presence of *S. pyrenaicus* in southern Iberia may be the result of recent colonization associated with river capture, as demonstrated our biogeographical reconstruction.

Keywords: Ancestral area reconstruction, Cenozoic, Phylogeny, Southern Iberia, Molecular clock, *Squalius*.

INTRODUCTION

The Betic Mountains and its foreland basin, the Guadalquivir River, which together bound the Hercynian Iberian Massif in its south and southeastern area, constitute the southern area of the Iberian Peninsula since the Miocene period. The Betics are part of the Gibraltar Arc along with the Moroccan Rif, which is located in the western end of the Alpine belt (Miller *et al.*, 2013). The southern area of the Iberian Peninsula constitutes a differentiated biogeographical region on the basis of freshwater fishes that encompasses not only the large Guadalquivir Basin but also several smaller independent river drainages in southwestern Portugal and in areas closet o the Gibraltar Strait in Spain (Figs. 1 y 2) (Doadrio, 1988; Filipe *et al.*, 2009). Several significant tectonic processes have formed the complex configuration of the hydrological network of southern Iberia, directly affecting the resident freshwater fauna, including the uplifting of the Betic Mountains in the Serravalian and Tortonian periods and events following the opening of the Gibraltar Strait at the Miocene-Pliocene boundary (Weijermans, 1991; Krijgsman *et al.*, 1999; Braga *et al.*, 2003), which have originated complex processes of speciation in a relatively small geographic region.

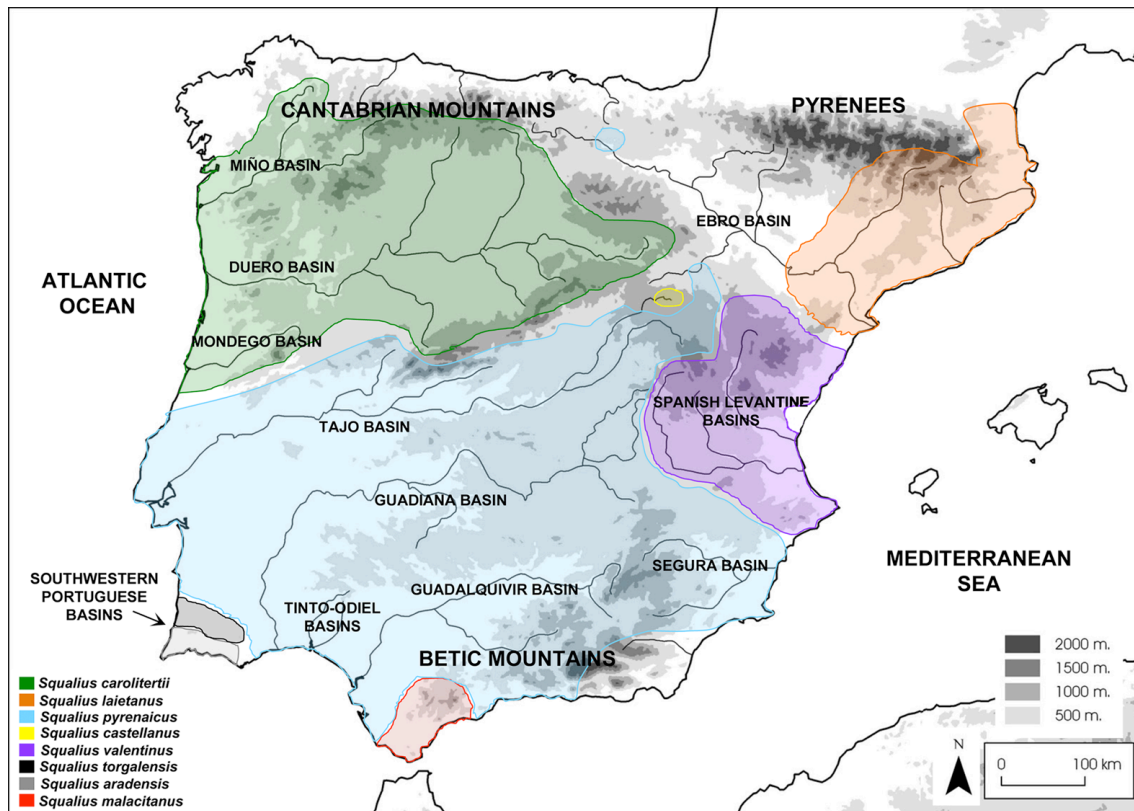


Fig. 1. Distribution range of Iberian *Squalius* species and location of Betic Mountains.

In addition to these macroevolutionary processes, more recent events could have influenced the evolutionary history of freshwater fishes in southern Iberia. The independent smaller basins have been affected since the Middle-Upper Pliocene by a heterogeneous annual hydrological cycle typical of the Circum-Mediterranean region and by catastrophic and erratic events (Mesquita *et al.*, 2005; Benito *et al.*, 2008; Lario *et al.*, 2011). In this fluctuating systems, the drying up of large river sections in summer often results in a series of isolated ponds, leading to high fish mortality and subsequent population bottlenecks (Coelho *et al.*, 1995; Magalhães *et al.*, 2002). In addition, sea level fluctuations during Pleistocene glaciations may have impacted habitat of aquatic species and played an important role in the population structure and genetic diversity of freshwater ichthyofauna (Gante *et al.*, 2009). Besides, current research on the Gibraltar Strait, supported by geological and geomorphological evidence, indicates that at least five tsunamis generated by strong earthquakes have affected this area during the past 7000 years, at a recurrence interval of 1200-1500 years (Lario *et al.*, 2011). These tsumanis may have been catastrophic for primary freshwater fishes from small rivers in southern Iberia. The genetic structure and population diversity of fishes inhabiting this area also reflect these recent climatic events. These microevolutionary processes related to climatic factors must have also contribute to shape the speciation pattern followed by freshwater organisms through the presence or absence of genetic exchange. The freshwater ecosystems of southern Iberia around the Gibraltar Strait are considered together a peninsular glacial refuge (Gómez and Lunt, 2007; Gante *et al.*, 2009).

To evaluate such events and their influence in the phylogenetic structure and speciation processes of aquatic communities, primary freshwater fish have been widely used as models (Doadrio *et al.*, 2002a,b; Perdices *et al.*, 2003; Böhlen *et al.*, 2006; Lopes-Cuhna *et al.*, 2012), since distribution of their lineages tends to reflect the history of river drainage systems rather than contemporary dispersal (Bernatchez and Wilson, 1998; Avise, 2000). The current hydrological network in the Iberian Peninsula was formed in the Upper Pliocene-Pleistocene periods (Calvo *et al.*, 1993; Pais *et al.*, 2012), and the general absence of connections among rivers has maintained the distribution of native fish, except in cases of human transfer, leading to vicariant speciation (Doadrio, 1988; Filipe *et al.*, 2009).

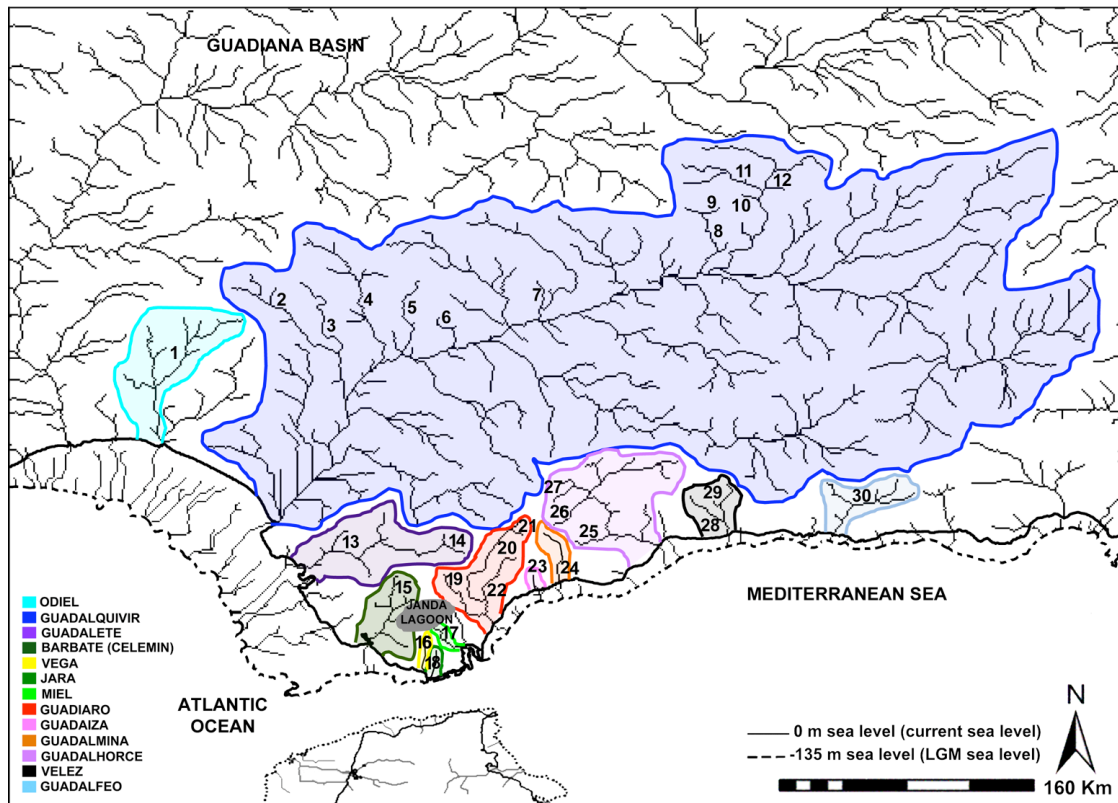


Fig. 2. Sampled river basins and bathymetric contours for the LGM and present-day sea levels.

One of the primary freshwater fish lineages currently inhabiting the Iberian Peninsula is *Squalius* Bonaparte, 1837. It is distributed throughout European freshwater lakes and rivers, and shows a high level of diversity in the Mediterranean area (Doadrio and Carmona, 1998; 2003; Zardoya and Doadrio, 1998; Durand *et al.*, 2000). Two main lineages, which diverged in the Miocene, are present in Europe: a group of large species, mainly occurring in central and northern Europe, and a second clade of small species present in the Mediterranean peninsulas with limited distribution ranges (Sanjur *et al.*, 2003; Perea *et al.*, 2010). In the Iberian Peninsula there are nine species of the genus *Squalius*, all of them endemic from different drainages systems and showing allopatric distribution ranges except for *S. alburnoides*, which co-occur with other species of *Squalius* in different basins due to their complex reproduction modes (Doadrio *et al.*, 2011; Fig. 1). In southern Iberia, Guadalquivir Basin and the smaller independent basins around the Gibraltar Strait, *Squalius* is represented by two species, with a vicariant pattern, of the Mediterranean lineage (*sensu* Sanjur *et al.*, 2003), *Squalius malacitanus* (Málaga chub) and *Squalius pyrenaicus* (Iberian chub). *Squalius malacitanus* is restricted to small rivers of southern Spain around the Gibraltar Strait (Doadrio and Carmona, 2006; Doadrio *et al.*, 2011). The distribution range of *Squalius pyrenaicus* covers mainly the southern half of the Iberian Peninsula, in particular the Tajo, Guadiana, Sado, Guadalquivir, Guadalete and Odiel basins, flowing into the

Atlantic Ocean, and the Segura, Guadalhorce, Velez, Guadalfeo and occasionally Ebro, river basins flowing into the Mediterranean Sea, being the Guadalhorce, Vélez and Guadalfeo basins the southern part of its range (Figs. 1 y 2). These southernmost Mediterranean populations along with the Atlantic Guadiana, Guadalquivir, Guadalete and Odiel basins have been considered as a monophyletic lineage relative to northern basins of Tajo, Sado and Ebro (Doadrio and Carmona, 2003).

It is assumed that both *S. malacitanus* and *S. pyrenaicus*, as a consequence of their status as primary freshwater fishes, should show clear phylogenetic and phylogeographical structure. For these reason they are suitable candidates for investigation of the effects of the complex tectonic and climatic changes occurred during the end of the Tertiary and Quaternary eras in the southern Iberian biogeographical region on their speciation processes. Similar to in other Iberian cyprinids with distribution across several basins, a strong inter- and intraspecific phylogenetic structure is expected, associated with geographical events (Doadrio *et al.*, 2002a,b; Doadrio and Carmona, 2003, 2004; Sousa-Santos *et al.*, 2007; Sousa *et al.*, 2008; Gante *et al.*, 2009). However, atypical or erratic events, such as Mediterranean climate weather patterns and the recent tsunamis occurring around the Gibraltar Strait, may led to more complex phylogenetic and phylogeographical patterns in this southern Iberian region. We hypothesize that the southern Iberian populations of the genus *Squalius* are sensitive to tectonic and climatic changes and that these changes are reflected in their current phylogenetic structure and genetic diversity. We suggest that larger drainage basins, particularly the Guadalquivir, should retain ancestral polymorphisms, whereas peripheral populations, mainly the ones of *S. malacitanus* because of its restricted distribution range, will present lower genetic diversity and marginal haplotypes, due to high mortality resulting from catastrophic Cenozoic events. Indeed, the range of *S. malacitanus*, an Endangered species (Freyhof and Kottelat, 2008), may be drastically reduced in the future as a result of global climate change and the trend to an increasingly arid Mediterranean region (Seager *et al.*, 2014).

The goals of this study were to use mitochondrial and nuclear DNA sequence data to (1) investigate phylogenetic patterns of *Squalius* species inhabiting southern Iberia; (2) evaluate the influence of Cenozoic tectonic and climatic changes on population genetic structure and diversity in those southern Iberian *Squalius* populations; (3) estimate divergence times among southern Iberian *Squalius* and (4) perform ancestral area reconstruction analysis of southern Iberian *Squalius* populations in order to unravel the evolutionary history of this genus in southern Iberia.

MATERIAL AND METHODS

1. Tissue and data collection

A total of 510 specimens of *Squalius malacitanus* y *S. pyrenaicus* from 13 drainage basins of southern Iberia were analyzed (Table 1; Fig. 2). Samples of *S. malacitanus* were collected from nearly its complete distribution range. Samples of *S. pyrenaicus* were mainly obtained from the southern distribution of the species in the Iberian Peninsula. Other Iberian populations of *S. pyrenaicus* as well as some individuals of all Iberian *Squalius* species but *S. albunoides* were included in the analyses in order to resolve the phylogenetic position of southern Iberian populations of the genus *Squalius* (Table 1; Fig. 2). For phylogenetic purposes of independent genes (Fig. S1), *S. squalus* and *S. laietanus*, belonging to the Euroasiatic lineage *sensu* Sanjur *et al.* (2003), were included as outgroups based on previous phylogenetic studies (Doadrio and Carmona, 2003, 2006; Perea *et al.*, 2010). Fish were captured by electrofishing under local authority permission, fin clipped, and returned to the stream, with the exception of 2-3 specimens per locality that were preserved in formalin for morphological identification. Fin clips were preserved in 95% ethanol. All voucher samples were stored in the Fish Collection and at the DNA Collection of the Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain.

2. DNA protocols

The complete mtDNA cytochrome *b* gene (*MT-CYB*; 1140 bp) of 510 individuals and a fragment of 1041 bp of the third exon of the nuclear Recombination Activating Gene 1 (*RAG1*) of 140 individuals, representing all southern *Squalius* populations, were amplified (Table 1). Primers and PCR protocols for both genes followed Perea *et al.*, (2010). After checking PCR products on 1% agarose gels, they were purified by ExoSAP-IT (USB, Cleaveland, OH, USA) and directly sequenced on Macrogen Inc (Amsterdam, Holanda; <http://www.macrogen.com>). All new sequences (haplotypes and alleles) were deposited in the GenBank database (Accession Numbers: *MT-CYB*: KU571153-KU571657; *RAG1*: KU571658-KU571692).

Sequences were aligned using the default pairwise and multiple alignment parameters in Clustal W (Thompson *et al.*, 1994) implemented in MEGA v.6.0 software (Tamura *et al.*, 2013). Nuclear data was phased with the PHASE algorithm (Stephens *et al.*, 2001; Stephens and Donnelly, 2003) in DNAsp v.5.0 (Librado and Rozas, 2009) using a probability threshold of 0.9 to resolve alleles. Recombination of the nuclear *RAG1* gene ($p=0.55$) was tested using the PHI test in SplitsTrees v.4.1.3. (Huson and Bryant, 2006).

Table 1. Data of analyzed *S. malacitanus* and *S. pyrenaicus* populations and GenBank Accession Numbers. Sample size refers to number of specimens analyzed for *MT-CYB/RAG1* genes.

Map site	Sample size (<i>MT-CYB</i> / <i>RAG1</i>)	River/Basin	Latitude / Longitud	GenBank Accession Numbers
<i>Squalius pyrenaicus</i>				
1	8/1	Odiel / Odiel	36.69 / -6.66	KU571653-KU571654 (<i>MT-CYB</i>)
2	3/0	Rivera de Huelva / Guadalquivir	37.95 / -6.62	-
3	36/3	Rivera de Cala / Guadalquivir	37.91 / -6.23	KU571617-KU571618; KU571630-KU571639 (<i>MT-CYB</i>)
4	3/0	Viar / Guadalquivir	38.09 / -6.14	KU571640-KU571642 (<i>MT-CYB</i>)
5	9/0	Huénar / Guadalquivir	37.86 / -5.74	KU571626-KU571629; KU571644 (<i>MT-CYB</i>)
6	9/2	Rivera de Ciudadreja / Guadalquivir	37.90 / -5.42	KU571621-KU571623; KU571625 (<i>MT-CYB</i>); KU571692 (<i>RAG1</i>)
7	4/3	Cabrillas / Guadalquivir	37.98 / -5.06	KU571643; KU571645 (<i>MT-CYB</i>)
8	6/4	Yeguas / Guadalquivir	38.32 / -4.24	KU571646-KU571649 (<i>MT-CYB</i>); KU571681-KU571684 (<i>RAG1</i>)
9	13/4	Cabrera / Guadalquivir	38.27 / -4.09	KU571685 (<i>RAG1</i>)
10	28/6	Robledillo / Guadalquivir	38.42 / -4.04	KU571619-KU571620; KU571624 (<i>MT-CYB</i>); KU571691 (<i>RAG1</i>)
11	9/4	Montoro / Guadalquivir	38.50 / -4.29	KU571650-KU571652 (<i>MT-CYB</i>); KU571686-KU571690 (<i>RAG1</i>)
12	4/0	Fresneda / Guadalquivir	38.57 / -3.77	KU571614-KU571616 (<i>MT-CYB</i>)
13	24/7	Guadalete / Guadalete	36.84 / -5.39	KU571608-KU571610 (<i>MT-CYB</i>)
14	4/3	Guadalete tributary / Guadalete	36.92 / -5.54	KU571606-KU571607 (<i>MT-CYB</i>); KU571679-KU571680 (<i>RAG1</i>)
25	26/0	Grande / Guadalhorce	36.73 / -4.86	KU571602-KU571605 (<i>MT-CYB</i>)
26	8/0	Pereilas / Guadalhorce	36.66 / -4.76	KU571599 (<i>MT-CYB</i>)
27	23/6	Turón / Guadalhorce	36.79 / -4.95	KU571600-KU571601 (<i>MT-CYB</i>); KU571677-KU571678 (<i>RAG1</i>)
28	30/7	Cuevas / Vélez	36.91 / -4.29	KU571595-KU571596 (<i>MT-CYB</i>); KU571675-KU571676 (<i>RAG1</i>)

29	19/6	Sabar / Vélez	36.94 / -4.23	KU571597-KU571598 (MT-CYB)
30	3/1	Guadaféo	36.88 / -3.42	KU571611-KU571613 (MT-CYB); KU571674 (RAG1)
-	8/0	Guadalquivir Basin	-	-
-	30/0	Guadiana Basin	-	-
<i>Squalius malacitanus</i>				
15	23/7	Celemin / Barbate	36.34 / -5.81	KU571580-KU571590 (MT-CYB); KU571670-KU571671 (RAG1)
16	20/10	Vega / Vega	36.07 / -5.66	KU571579 (MT-CYB); KU571666-KU571669 (RAG1)
17	18/6	Miel / Miel	36.13 / -5.45	KU571567-KU571571 (MT-CYB); KU571662-KU571663 (RAG1)
18	16/8	Jara / Jara	36.01 / -5.60	KU571572-KU571578 (MT-CYB); KU571664-KU571665 (RAG1)
19	21/7	Hozgarganta / Guadiaro	36.43 / -5.45	KU571557-KU571558 (MT-CYB)
20	15/6	Guadiaro / Guadiaro	36.65 / -5.28	KU571559 (MT-CYB); KU571658 (RAG1)
21	18/8	Guadalevín / Guadiaro	36.74 / -5.17	KU571553-KU571556 (MT-CYB); KU571672-KU571673 (RAG1)
22	35/7	Genal / Guadiaro	36.56 / -5.21	KU571560-KU571566 (MT-CYB)
23	9/5	Guadaiza / Guadaiza	37.91 / -6.23	KU571661 (MT-CYB)
24	33/6	Guadalmina / Guadalmina	36.52 / -5.05	KU571591-KU571594 (MT-CYB); KU571659-KU571660 (RAG1)
Other <i>Squalius</i> species				
<i>Squalius aradensis</i>			AJ698711, AJ852502 (MT-CYB)	
<i>Squalius castellanus</i>			DQ521423 (MT-CYB); HM560575 (RAG1)	
<i>Squalius carolitertii</i>			AJ698453-AJ698456 (MT-CYB); HM560446 (RAG1)	
<i>Squalius laietanus</i>			HM560187-HM560188 (MT-CYB)	
<i>Squalius torgalensis</i>			HM560204-HM560205 (MT-CYB)	
<i>Squalius valentinus</i>			KR871719-KR871767 (MT-CYB); KR871768-KR871795 (RAG1)	

3. Phylogeny and population structure

For phylogenetic analysis of independent genes, sequences of *MT-CYB* and *RAG1* were collapsed to haplotypes using the web-based program ALTER (González-Peña *et al.*, 2010). For concatenated phylogeny, the dataset including specimens amplified for both genes was used. The best-fit model of molecular evolution for each gene and codon position was estimated via AIC in Par PartitionFinder v.1.1.1 (Lanfear *et al.*, 2012). Best partition schemes used in different phylogenetic analyses based on the selected evolutionary models are presented in Table S1 (Supporting Information of this chapter). Bayesian Inference (BI) was performed using MrBayes v.3.2 (Ronquist *et al.*, 2012). Two simultaneous analyses were run for 10^7 generations, each with four MCMC chains sampling every 100 generations. Convergence was checked in Tracer v.1.6 (Rambaut and Drummond, 2007). After discarding the first 10% of generations as burn-in, the 50% majority rule consensus tree and posterior probabilities were obtained. Phylogenetic analyses of both genes, independently and concatenated, were conducted using Maximum Likelihood (ML) in RaxML software implemented in the Trex-online server (Stamakis, 2006) using the substitution model GTRGAMMAI and the rapid bootstrap algorithm (Stamakis *et al.*, 2007) for both *MT-CYB* and concatenated analyses. Node confidence in ML analyses was estimated by rapid bootstrapping using a random seed (1000 replicates). Uncorrected-*p* distances for both genes were estimated in MEGA v.6.0. (Tamura *et al.*, 2013) in order to quantify genetic differences between species and among their populations. To assess the phylogeographic history of *Squalius* in southern Iberia, we reconstructed two independent networks for *MT-CYB* y *RAG1* using HaploViewer (available in <http://www.cibiv.at/~%20greg/haploviewer>).

We estimated levels of genetic diversity based on *MT-CYB*, including number of haplotypes (*h*), haplotype diversity (H_D) and nucleotide diversity (π), using DNAsp v.5.0 (Librado and Rozas, 2009) for southern Iberian populations of the genus *Squalius*. To assess the relative contribution of genetic variation to structure within and between populations in *S. malacitanus* and southern populations of *S. pyrenaicus*, we performed several analyses of molecular variance (AMOVA) implemented in Arlequin v.3.5.1.2 (Excoffier and Lischer, 2010), testing for genetic structure between groups of populations (Atlantic and Mediterranean) and among basins in both species. Global genetic structure of *S. malacitanus* and *S. pyrenaicus* considering their geographical location was also investigated using a Spatial Analysis of Molecular Variance (SAMOVA; Dupaunloup *et al.*, 2002). We run SAMOVA from *k*=2 to *k*=13 (all southern Iberian basins), with 500 and 1000 simulated annealing processes. Each run was repeated three times to check consistency. The genetic differentiation among populations was addressed through *MT-CYB* F_{ST} pairwise comparisons (Hudson *et al.*, 1992) and estimates of gene flow as virtual number

of migrants (N_m) exchanged among populations per generation (Slatkin and Barton, 1989), using Arlequin (Excoffier and Lischer, 2010). Significant deviations from the null hypothesis of no differentiation were assessed with 10000 permutation tests. As multiple paired tests were performed, p -values were adjusted by Bonferroni's correction (Rice, 1989).

4. Species tree and divergence time estimates

A multilocus species tree analysis of Iberian populations of the genus *Squalius* belonging to the phylogenetic clade *S. malacitanus* + *S. pyrenaicus* + *S. valentinus* (Doadrio and Carmona, 2006; Perea et al., 2010; results from this study) was carried out using the concatenated matrix (*MT-CYB* + *RAG1*) and the multispecies coalescent method *BEAST implemented in BEAST v.1.7 (Heled and Drummond, 2010; Drummond et al., 2012). This analysis allows incorporating information on population to the species tree analyses through the coalescent theory and requires predefined groups to be tested, therefore, we considered the monophyletic clades obtained in previous performed phylogenetic and network analyses as different clusters. This analysis was also used to estimate divergence time among southern Iberian *Squalius* using a lognormal relaxed molecular clock and a Yule speciation model. We calibrated the molecular clock, incorporating as normal prior the evolutionary rate of 0.4% divergence per million years for the *MT-CYB* gene proposed for cyprinids of the subfamily Leuciscinae, which is based on fossil calibrations (Perea et al., 2010) and estimated the evolutionary rate of the nuclear *RAG1* gene relative to the *MT-CYB*. MCMC analyses were run for 50 million generations, with parameters logged every 10000 generations. The remaining parameters were used as default. Output from *BEAST was examined in TRACER v.1.6 (Rambaut and Drummond, 2007), and the results were summarized using TreeAnnotator v.1.7 (Drummond et al., 2012).

5. Ancestral area reconstruction

We reconstructed the ancestral states of the distribution ranges of the southern Iberian populations of the genus *Squalius* using the Dispersal-Extinction-Cladogenesis (DEC) model of LAGRANGE (Ree et al., 2005; Ree and Smith, 2008) and the statistical DIVA (S-DIVA; Yu et al., 2010) implemented in RASP 3.1 (Yu et al., 2015). The ultrametric and dichotomous tree obtained for the *MT-CYB* gene in *BEAST analysis (containing all populations of southern Iberian *Squalius*) was used as the tree topology on which mapping ancestral areas. Tips of the ultrametric tree were code by areas (basins) in order to reconstruct ancestral states.

6. Sea level patterns

We employed a geographic information system to visualize changes in sea level from the LGM, when sea level was 135 m lower than currently (Miller *et al.*, 2005), to the present. Datasets used to generate a sea level change map were obtained from the Digital Chart of the World (ESRI, 1993) and manipulated in ArcMap v.10 (ESRI, Redlands, CA). Bathymetric data were obtained from a 30 arc-second (ca. 1 km) dataset, GEBCO 08 (www.gebco.org; September 2010 release) and manipulated to produce bathymetric contours and sea-level drainage patterns using the hydrological tools in ArcInfo.

RESULTS

1. Phylogeny and population structure

The phylogenetic analysis based on mitochondrial (*MT-CYB*; 1140 bp) and concatenated (*MT-CYB* + *RAG1*; 2181 bp) genes and both phylogenetic methods performed (ML and BI) were congruent and relationships among Iberian *Squalius* species belonging to the Mediterranean lineage (Sanjurjo *et al.*, 2003) were highly supported and consistent with published studies (Doadrio and Carmona, 2006; Almada and Sousa-Santos, 2010; Perea *et al.*, 2010). All Iberian species of the genus *Squalius* were monophyletic. From the root of the phylogenetic tree to the tips, southwestern Portuguese species of the genus, *S. aradensis* and *S. torgalensis*, appeared as the sister group of the rest of Iberian species; the second divergent clade was that constituted by *S. carolitertii* and *S. castellanus*, and finally, *S. malacitanus* was the sister group of a monophyletic clade formed by *S. pyrenaicus* and *S. valentinus* (Figs. 3 and S1).

With regards to southern Iberian populations of the genus *Squalius*, BI and ML analyses of *MT-CYB* and *MT-CYB* + *RAG1* revealed two main clades: one containing samples attributed to *S. malacitanus* and the other to *S. pyrenaicus*, being this later group sister to the Spanish Levantine species *S. valentinus* (Figs. 3 and S1). Nuclear phylogeny also separated *S. malacitanus* and *S. pyrenaicus* (Fig. S2). Mitochondrial and nuclear phylogenies supported the structure of *S. malacitanus* into two subclades (Figs. 3 and S1), one comprising the Mediterranean populations (Guadiaro, Guadalmina and Guadaiza basins) and the other Atlantic populations from the Celemin (Barbate Basin), Jara, Vega and Miel river basins. Miel Basin is located near the Gibraltar Strait and currently drains into the Mediterranean Sea as its junction with the Atlantic Ocean. Accordingly, although this river flows into the Mediterranean slope, we assigned its populations to the Atlantic, due to its geographic proximity to, and common history

with, those populations (Fig. 2). *Squalius pyrenaicus* did not present a geographically well-supported phylogenetic structure, and Mediterranean populations (Guadalhorce, Vélez and Guadalfeo basins) clustered in the phylogenetic tree together with haplotypes observed in some tributaries of the middle part of the Guadalquivir Basin and a tributary of the Upper Guadalete Basin on the Atlantic slope (Figs. 3 and S1).

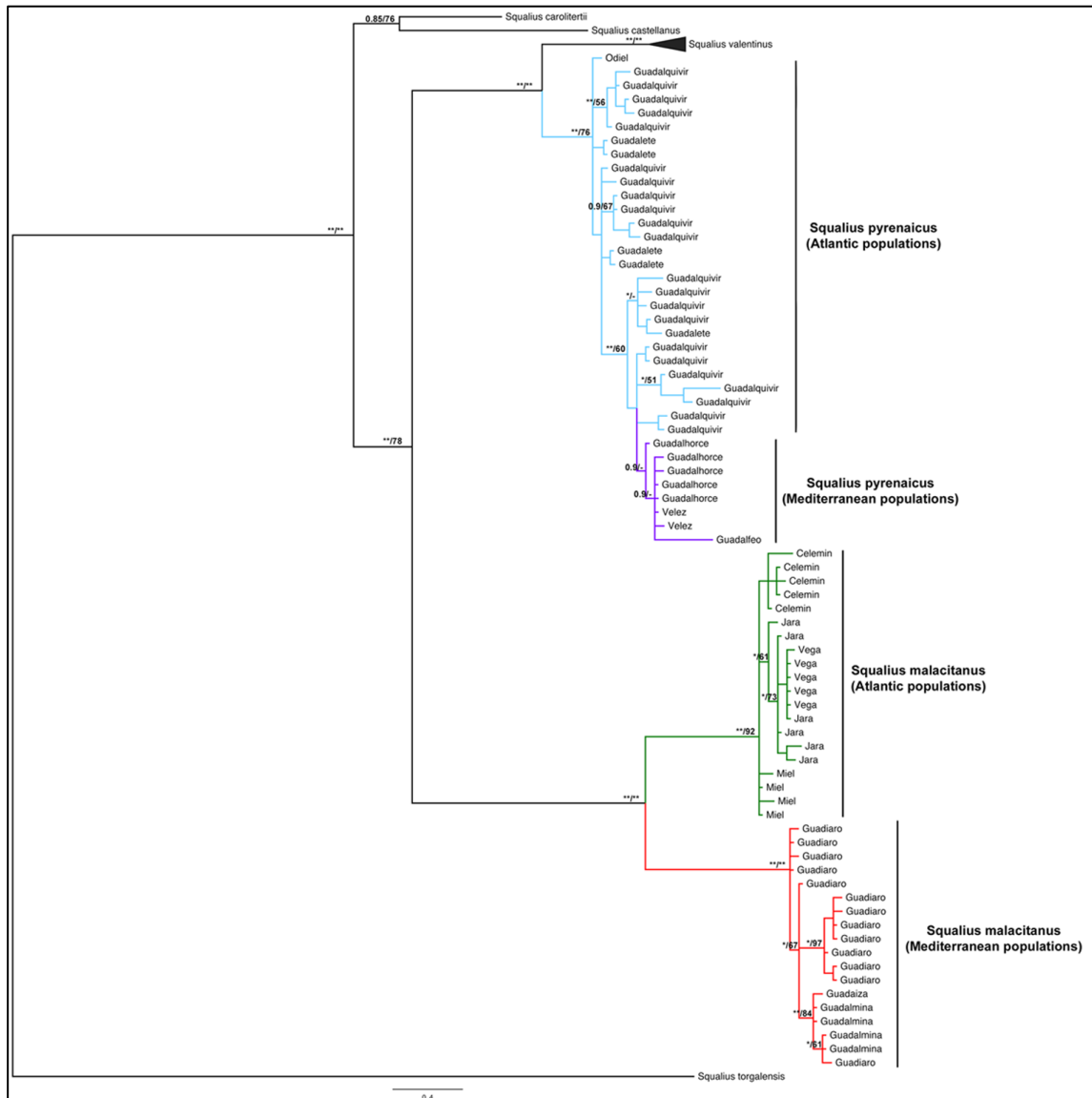


Fig. 3. Phylogenetic tree of *S. malacitanus* and *S. pyrenaicus* populations analyzed in this study derived from the BI and ML analyses of the concatenated dataset (MT-CYB + RAG1). Posterior probability values for BI and bootstrap values for ML (1000 replicates) are shown before and after slash respectively. Labels indicate river system. ** means posterior probability = 1. * means posterior probability = 0.95.

Overall uncorrected-*p* genetic distances for the *MT-CYB* gene varied from 0.01% observed in specific pairwise comparisons of Mediterranean populations of *S. pyrenaicus* to 8.4% between Mediterranean populations of the two species (Table S2 in the Supporting Information of this chapter). Within *S. malacitanus*, genetic distances between Atlantic and Mediterranean populations for the *MT-CYB* gene ranged from 4.1% to 4.5%, while this range was 0.6-1% between Atlantic and Mediterranean basins in southern Iberian populations of *S. pyrenaicus*. Uncorrected-*p* genetic distances for the nuclear *RAG1* gene ranged from 0.4 to 0.8% between *S. malacitanus* and southern Iberian populations of *S. pyrenaicus*, from 0.1% to 0.2% in *S. malacitanus* and from 0.1% to 0.4% in *S. pyrenaicus* (Table S2 in the Supporting Information of this chapter).

Mitochondrial and nuclear network analyses of both *Squalius* species (Fig. 4) matched the well-supported lineages recovered in the phylogenetic tree: one group for *S. pyrenaicus* and two groups for *S. malacitanus* (Mediterranean and Atlantic populations). In *S. malacitanus*, no mitochondrial haplotype or nuclear allele was shared between the Atlantic (Jara, Vega, Miel and Celemín) and Mediterranean (Guadiaro, Guadalmina and Guadaiza) populations, demonstrating high genetic and geographic separation. In *S. pyrenaicus*, no mitochondrial haplotypes were shared between Atlantic (Guadalquivir and Guadalete) and Mediterranean (Guadalhorce, Vélez and Guadalfeo) basins, however, sharing of nuclear alleles among populations was common (Fig. 4).

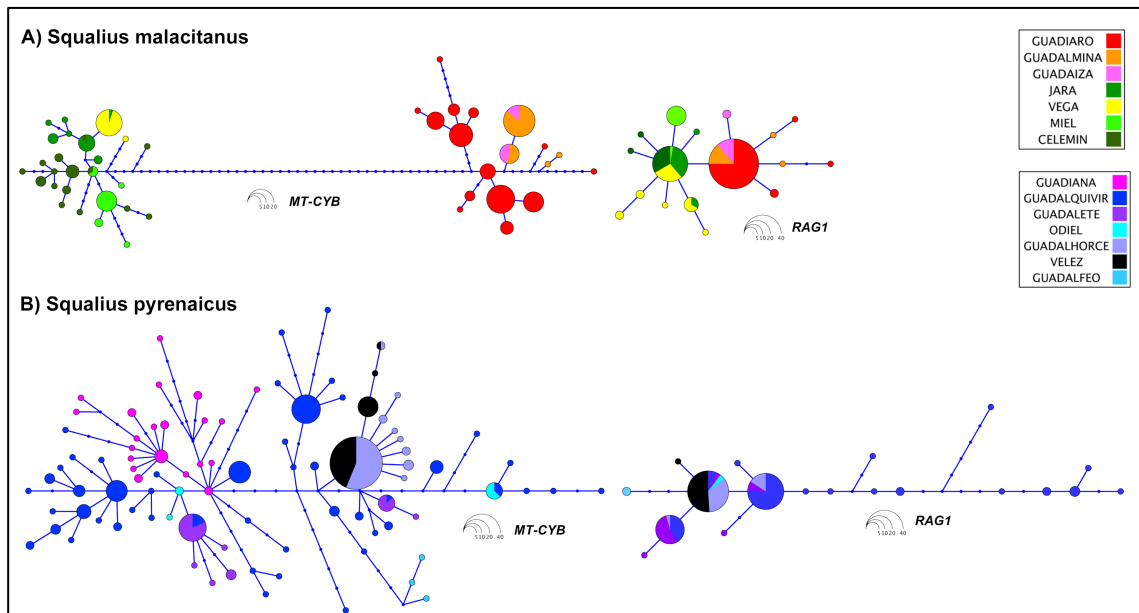


Fig. 4. Mitochondrial and nuclear DNA haplotype networks. The size of the circles is proportional to the frequencies of the haplotypes. Colors represent river systems. Number of nodes are the mutational steps.

Genetic diversity values were lower for analyzed *S. pyrenaicus* populations than for *S. malacitanus* (Table S3 in the Supporting Information of this chapter). Haplotype and nucleotide diversity of Atlantic populations of *S. malacitanus* was higher than in Mediterranean populations. Within Mediterranean populations, values of genetic diversity were high only in the Guadiaro Basin (Table S3 in the Supporting Information of this chapter). *Squalius pyrenaicus* also showed a similar pattern of higher genetic diversity values in Atlantic populations relative to Mediterranean populations; these values were higher in the larger basin, the Guadalquivir (Table S3 in the Supporting Information of this chapter).

AMOVA analyses supported high and significant genetic differentiation among and within populations (Atlantic vs. Mediterranean) of both species (Table S4 in the Supporting Information of this chapter). In *S. malacitanus*, higher percentages of genetic differentiation were found among basins than within basins. In *S. pyrenaicus*, the within-basin genetic variation was higher than among basins, especially in the larger (Guadalquivir). SAMOVA analysis of both species revealed the highest differentiation among groups and the smallest percentage of variation within groups for $k=4$, corresponding these four groups to Atlantic and Mediterranean populations in both species, which proved to be the best partition of genetic variation hypothesis according to geographic criteria (Table 2).

Table 2. SAMOVA analysis of MT-CYB data for $k=4$ (Atlantic and Mediterranean populations of *S. malacitanus* and Atlantic and Mediterranean populations of *S. pyrenaicus*). Df = degrees of freedom; SSD = sum of squared deviation.

Source of variation	df	SSD	Variance components	Percentage of variation
Among groups	3	7625.760	34.23596	95.17
Among populations				
Within groups	13	227.095	0.93613	2.60
Within populations	299	239.458	0.80086	2.23
Total	315	8092.313	35.97295	

Pairwise F_{ST} values are presented in Table S5 (Supporting Information of this chapter). In both species, the lowest differentiation values were found between basins belonging to the same slope, whereas the highest pairwise values involved basins on different slopes. An exception was the Miel Basin, geographically belonging to the Mediterranean slope but showing lower F_{ST} values in pairwise comparisons that placed this basin with Atlantic populations. These results of genetic differentiation were congruent with the restricted or absent gene flow found between Atlantic and

Mediterranean populations of both species, as was reflected by N_m values lower than 1 (Table S5 in the Supporting Information of this chapter). Nevertheless, evidence of gene flow ($N_m > 1$) was found among rivers of the same slope in both species, especially among Mediterranean populations, although gene flow was also significant among Atlantic populations of *S. pyrenaicus*.

2. Species tree and divergence time estimations

Species tree analysis using *BEAST and both *MT-CYB* and *RAG1* genes recovered with high support (pp = 1) Atlantic and Mediterranean populations of *Squalius malacitanus* as different groups. The split between *S. malacitanus* and the clade comprised of *S. pyrenaicus* + *S. valentinus* was estimated to take place 11.5 Ma (6.8-18.6 Ma), at the boundary of the Serravalian and Tortonian periods. Nevertheless, although posterior probability isolating Atlantic and Mediterranean populations in *S. pyrenaicus* was also high (pp=1) divergence between these two groups of populations was very recent (Fig. 5). The divergence of *S. pyrenaicus* from its sister species *S. valentinus* was estimated to have occurred during the Middle Pliocene 3.9 Ma (1.3-8.1 Ma). The separation between the Mediterranean and Atlantic populations of *S. malacitanus* was estimated to be 4.9 Ma (2.1-8.5 Ma), closet o the Miocene-Pliocene boundary, while the isolation of the Atlantic and Mediterranean populations of *S. pyrenaicus* took place at 0.6 Ma (0.3-1.0 Ma), in the Pleistocene period (Fig. 5).

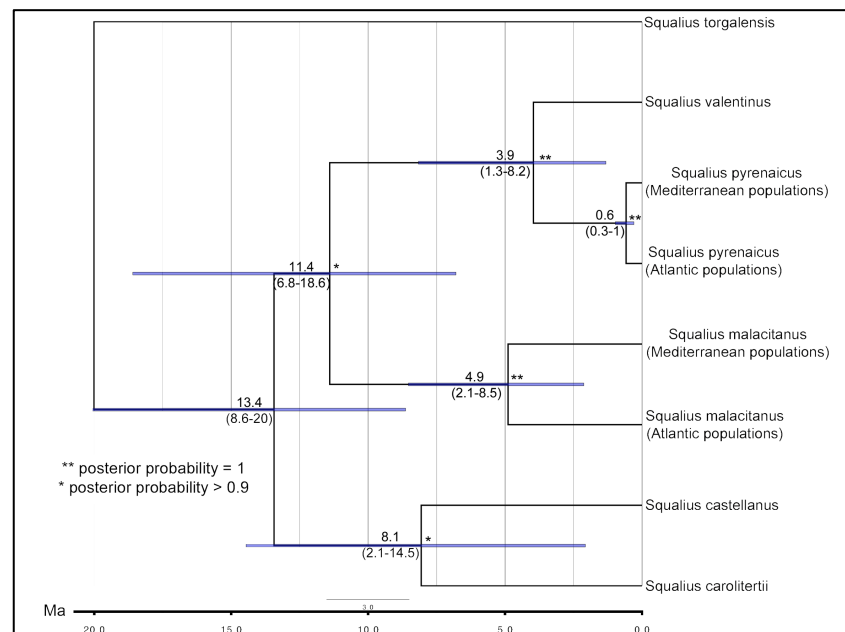


Fig. 5. Species tree and divergence times for *S. malacitanus* and southern Iberian populations of *S. pyrenaicus* estimated by BEAST software. Analysis is based on concatenated *MT-CYB* and *RAG1* genes, a relaxed molecular clock, and an evolutionary rate of 0.4% per lineage per million years. Bayesian probability >0.95 (*). Bayesian probability = 1 (). Above branches, divergence time estimates; below branches 95% HPD confidence intervals.**

3. Ancestral area reconstruction

Ancestral area reconstruction using LAGRANGE (DEC model) and S-DIVA methods showed a complex biogeographical history followed by southern Iberian populations belonging to both species, *S. malacitanus* and *S. pyrenaicus*, in which dispersal and vicariant events were frequent (Figs. 6 y 7). In general, DEC analyses estimated more dispersal events than S-DIVA for the whole phylogenetic tree and some of the ancestral nodes could not be reconstructed by these methods with high support, such is the case of the ancestral area explaining the separation of *S. malacitanus* from the *S. pyrenaicus* + *S. malacitanus* clade. Nonetheless, biogeographical routes followed by Atlantic and Mediterranean populations of both species were highly congruent between both geographical reconstructions, especially in those nodes highly supported by phylogenetic analyses. These geographical routes are explained below in the Discussion section.

4. Sea level change patterns

The bathymetric reconstruction showed a posible downstream connection only between Jara and Vega basins during the LGM. On the other hand, headwaters of most small basins on the Atlantic slope were in the proximity of the ancient basin of the Janda Lagoon (Fig. 2).

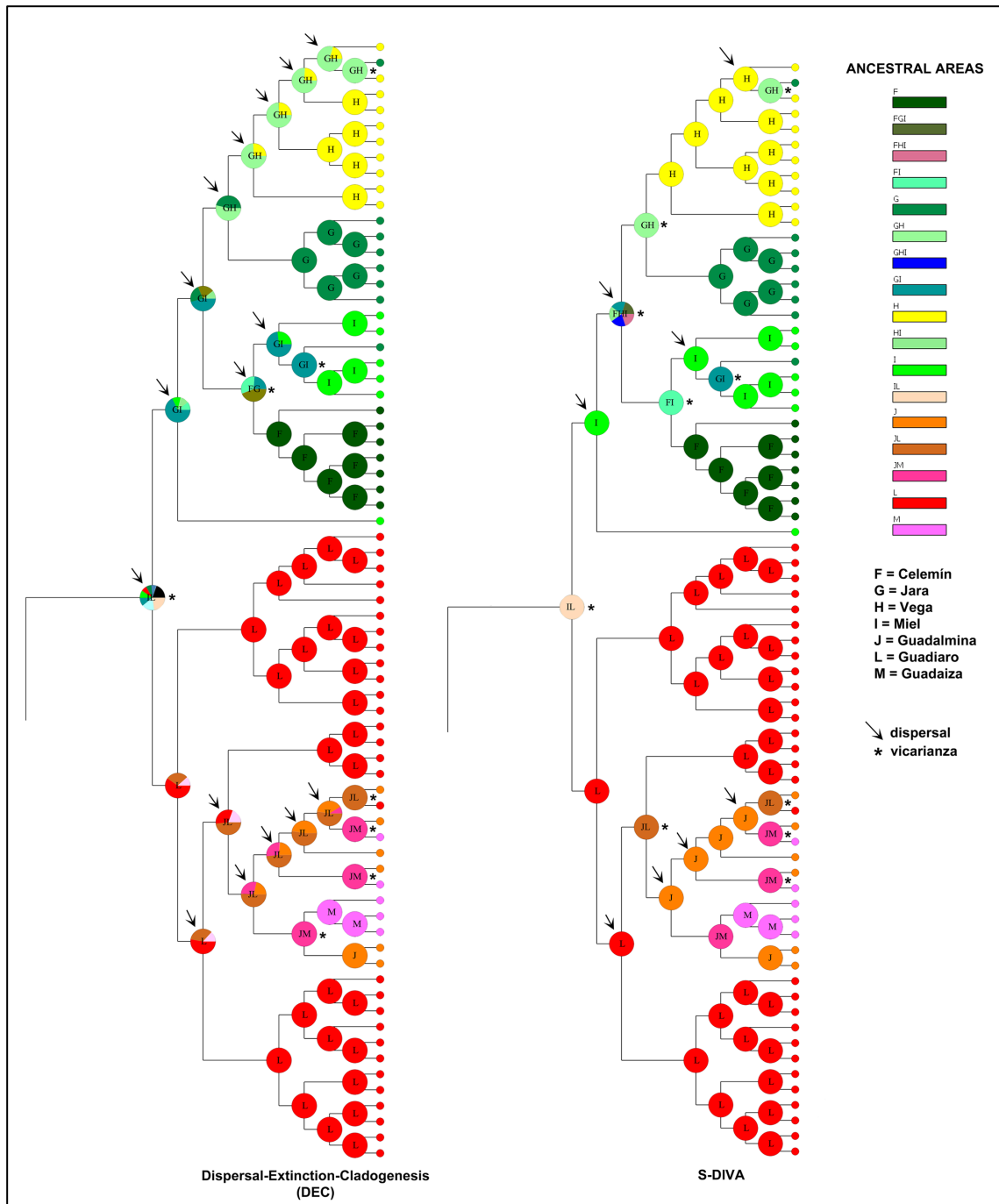


Fig. 6. Ancestral area reconstruction of *S. malacitanus* using the Dispersal-Extinction-Cladogenesis (LAGRANGE) and the Statistical-DIVA (S-DIVA) models implemented in RASP software. * denotes vicariant events. Arrows indicate dispersal events.

DISCUSSION

Southern Iberian Peninsula represents one of the most complex areas to carry out biogeographic and diversification studies due to early orogenic and tectonic events that led to diversification processes during Miocene, forming old endemic Mediterranean lineages in various taxonomic groups (Martínez-Solano *et al.*, 2004; Miraldo *et al.*, 2011). The Miocene separation of the European-African continental margin and the reopening of the Gibraltar Strait in the Miocene-Pliocene boundary played a key role in shaping genetic diversity and biogeography of Western Mediterranean species, including fishes (Machordom and Doadrio, 2001; Doadrio and Perdices, 2005).

1. Interspecific diversification of southern Iberian *Squalius*

We demonstrate that the evolutionary history of *S. malacitanus* and southern Iberian populations of *S. pyrenaicus*, the two *Squalius* species occurring in southern Iberia in the region close to the Gibraltar Strait, displayed similar geographical divergence patterns, including a primary separation between Atlantic and Mediterranean populations of both species, most conspicuously in *S. malacitanus*, due to the close relationship of Mediterranean populations of *S. pyrenaicus* with some of the tributaries of the middle-upper part of the main Atlantic basins (Guadalquivir and Guadalete). However, the time, and consequently, the level, of divergence of Atlantic from Mediterranean populations differ. Thus, it can be assumed that distinct geological or climatic events at different temporal scales were responsible for the phylogenetic and biogeographic patterns of Atlantic and Mediterranean populations of both species. The endemic local species, *S. malacitanus*, presented a deeper divergence of Atlantic from Mediterranean populations compared to southern populations of *S. pyrenaicus* close to Gibraltar Strait, probably as a consequence of its ancient origin in the southern Iberian area (Fig. 5).

Our molecular clock dates the separation between *S. malacitanus* and the clade grouping *S. pyrenaicus* and *S. valentinus* at ca. 11.5 Ma, closet o the Serravalian-Tortonian boundary in the Upper Miocene (11.6 Ma). Similar divergence times have been found in other Iberian and North African vertebrate species around the Gibraltar Strait (Carranza *et al.*, 2006; Carretero, 2008). For amphibians and reptiles inhabiting the southern Iberia, the Tortonian has been considered an important period for isolation and diversification, driven by seaways opening within the Betic and Rifian areas (Fromhage *et al.*, 2004; Albert *et al.*, 2007). Among freshwater fish species around the Gibraltar Strait, the geographical pattern and the temporal scale of the separation between *S. malacitanus* and the *S. pyrenaicus* + *S. valentinus* clade is unusual, since splitting of these species occurred between the Iberian area of the Gibraltar Strait and other

Iberian basins and not between the southern Iberian Peninsula and North Africa. In other known freshwater fish species, the most common divergence pattern is the vicariance between southern Iberian and northern Moroccan populations associated with the opening of the Gibraltar Strait in more recent times than our estimate (Doadrio, 1990; Perdices *et al.*, 1995, 2001; Machordom and Doadrio, 2001; Doadrio and Perdices, 2005; Almada *et al.*, 2009).

The estimated time of divergence of the analyzed *Squalius* species of southern Iberia predates the refilling of the Mediterranean Sea by the opening of Gibraltar Strait through the Zanclean flood (5.3 Ma), hence some older vicariant events must be involved (Esteban *et al.*, 1996; Krijgsman *et al.*, 1999; García-Castellanos *et al.*, 2009). Both ancestral area reconstructions were not able to resolve the state of the node separating *S. malacitanus* from the *S. valentinus* + *S. pyrenaicus* clade, although the most probable route in DEC analyses estimated a dispersal toward southern Iberia and a posterior vicariance among distribution ranges: however, marginal probability supporting this biogeographical route was low (DEC = 0.0005; S-DIVA = 0.125). The most plausible hypothesis to explain the divergence of *S. malacitanus* from the *S. valentinus* + *S. pyrenaicus* clade is its separation by the orogenic uplifting and tectonic movements related to the formation of the Betic Mountains during the Tortonian period and its associated foreland, the Guadalquivir Basin, along with the formation of the southern hydrological network as uplifting and tectonic movements occurred. This resulted in the emergence of island separated by seaways and the reconfiguration of the freshwater basins in the region when the seaways were closed, isolating *S. malacitanus* from other Iberian *Squalius* species and favoring its speciation process (Betzler *et al.*, 2006; Martin *et al.*, 2009, 2014).

In contrast to the limited distribution range of *S. malacitanus*, confined to rivers around the Gibraltar Strait, *S. pyrenaicus* is widespread throughout the central and southern Iberian Peninsula (Fig. 2). According to our molecular clock, *S. pyrenaicus* diverged from its sister species *S. valentinus* ca. 4 Ma during the Middle-Upper Pliocene, related to the transformation of the large Iberian Miocene endorheic lagoons into the current exoreic basins, which began in this period and culminated in the Pleistocene associated with tectonic reactivation of the region and resulting in tilting and local deformation of Neogene formations (Calvo *et al.*, 1993; Pais *et al.*, 2012). This timing of divergence and diversification during the Middle-Upper Pliocene is concordant with that reported for other *Squalius* species in the Iberian Peninsula, including *S. valentinus* (Doadrio and Carmona, 2006; Perea and Doadrio 2015), *S. castellanus* (Doadrio *et al.*, 2007b) and *S. laietanus* (Doadrio *et al.*, 2007a). The Plio-Pleistocene has been considered an important period of diversification for Iberian cyprinids in general (Zardoya and Doadrio, 1998; Robalo *et al.*, 2006; Sousa-Santos *et al.*, 2007, 2014; Almada and Sousa-Santos, 2010), así como para otra fauna de agua dulce (Ribera and Vogler, 2004; Ribera *et al.*, 2011). The

ancestral area reconstruction based on DEC and S-DIVA estimated a vicariance between the Spanish Levantine area, distribution range of *S. valentinus*, and the Guadalquivir Basin, in which inhabits *S. pyrenaicus*, as the most probable event isolating both species (marginal probability DEC=0.2; marginal probability S-DIVA=1.0; Fig. 7). This vicariant event may be associated to the abovementioned formation of Iberian exoheric basins (Calvo *et al.*, 1993; Pais *et al.*, 2012).

2. Intraspecific diversification of *Squalius malacitanus*

The species *Squalius malacitanus* showed intraspecific phylogenetic and geographical structure on the basis of the mitochondrial and nuclear genes analyzed. Therefore, all analyses performed in this study supported the subdivision of *S. malacitanus* into two groups, Atlantic and Mediterranean populations. Our molecular clock showed that separation of Atlantic from Mediterranean populations dated from ca. 5 Ma, close to the Miocene-Pliocene boundary. The same divergence pattern and timing of divergence between Atlantic (southwestern clade) and Mediterranean (central clade) populations is found in the reptile genus *Blanus* in southern Iberia (Albert *et al.*, 2007). Subsequent to the opening of the Gibraltar Strait during the Pliocene, the areas nearest the present-day coast of southern Iberia remained marine basins, but were progressively filled by Neogene sediment with the uplift of the Betic Mountains throughout this period (Braga *et al.*, 2003). This event would lead to the development of new freshwater habitats in the southernmost area of the Iberian Peninsula, close to the Gulf of Cádiz and Gibraltar Strait, and the progressive isolation of Atlantic populations of *S. malacitanus* from the Mediterranean populations. Indeed, both DEC and S-DIVA reconstructions estimated a common ancestral area shared by Guadiaro and Miel (Mediterranean slope) and Jara basins (Atlantic slope) as the most probable for both Atlantic and Mediterranean populations (Guadiaro) from genetically Atlantic populations (Miel and Jara) (marginal probability DEC=0.09; marginal probability S-DIVA=1; Fig. 6).

The high genetic diversity and low genetic structure found in *S. malacitanus* Atlantic populations, as demonstrated the predominance of dispersal events among these Atlantic basins in both DEC and S-DIVA reconstruction (Fig. 6), cannot be explained by downstream connections among basins except for Jara and Vega basins, as shown by our bathymetric reconstruction (Fig. 2), which shared a common ancestral area (marginal probability DEC = 0.4; marginal probability S-DIVA = 1.0; Fig. 6). A possible explanation for this genetic pattern (high genetic diversity and low genetic structure) is the presence of the Janda Lagoon during the Quaternary (Fig. 2). This lagoon, which in periods of heavy rainfall reached an area up to 50 Km², was drained in the mid-twentieth century for pasture and agriculture (Recio, 2007). This large lagoon is likely to have played an important role in genetic diversity and structure of *S. malacitanus* and other freshwater

fish species inhabiting the Iberian slope of Gibraltar Strait during the Pleistocene.

The Janda Lagoon, has a tectonic origin related to the collision of the European and African plates, but it was not until the Middle-Upper Pleistocene that it acquired the form of an endorheic freshwater lagoon fed by the Barbate Basin, from which Celemin River is its main tributary, and adjacent basins that did not flow into the sea. In the LGM period, the lagoon was connected to the sea during the Flandrian transgression 15,000 years ago, when the Barbate Basin was open to the sea. Fluvial sediments in the narrow strait that linked the lagoon with the sea blocked the lagoon-sea connection again 3800 year (Goy *et al.*, 1995; Luque *et al.*, 1999). The lagoon allowed connection and gene flow among Atlantic populations of species such as *S. malacitanus*, supporting the sharing of mitochondrial haplotypes (low genetic structure; Table S2 in the Supporting Information of this chapter and Fig. 4) and high genetic diversity (Table S3 in the Supporting Information of this chapter). The lagoon likely harbored high effective populations of freshwater fish population, acting as a reservoir of freshwater organisms.

Within Mediterranean populations of *S. malacitanus*, Guadiaro and Guadalmina basins did not share mitochondrial haplotypes, as evidenced the limited gene flow estimated between these two basins. Nevertheless, gene flow between Guadalmina and Guadaiza, two adjacent basins, was high (Table S5 in the Supporting Information of this chapter). DEC and S-DIVA reconstructions estimated a colonization Guadiaro to Guadalmina (marginal probability DEC = 0.26; marginal probability S-DIVA = 1.0; Fig. 6) and a second colonization from either an ancestral area shared by Guadiaro and Guadalmina to Guadaiza (marginal probability S-DIVA = 1.0; Fig. 6). These biogeographical relationships were also supported by the presence of one highly frequent nuclear allele shared by these three basins, which it has not been sort yet as consequence of the recent colonization of the Guadalmina and Guadaiza basins (Fig. 4).

Our bathymetric reconstruction showed that mouths of these Mediterranean basins were widely separated during the LGM (Fig. 2). The continental shelf at the Mediterranean slope of the Gibraltar Strait is extremely narrow as consequence of old deep canyons formed by regressive erosion caused by an east-flowing river during the time that the strait was opening (Blanc, 2002; Iagel and Van Den Driessche, 2006; García-Castellanos *et al.*, 2009). Therefore, continental shelf in this Mediterranean area has been similar in width since the Miocene-Pliocene boundary, not allowing downstream connections among river basins. Guadiaro Basin has an arched shape and headwaters of some of its tributaries are very close to the headwaters of Guadalmina and Guadaiza, being the area of connection among these basins (Fig. 2). In addition, these Mediterranean basins flow throughout mountainous areas before reaching the sea. This complex geomorphology promoted vicariant events after the colonization of the different basins of

southern Mediterranean region by *S. malacitanus* leading to the isolation of river basins and the presence of isolated and highly diverse mitochondrial lineages (Guadiaro and Guadalmina + Guadaiza), as well as the presence of low frequent private nuclear alleles in each of these Mediterranean populations of *S. malacitanus*.

3. Intraspecific diversification of southern Iberian populations of *S. pyrenaicus*

The divergence of Atlantic from Mediterranean populations of *S. pyrenaicus* is dated at the Pleistocene, ca. 0.6 Ma, when the hydrographic network of the Iberian Peninsula was well established (Calvo *et al.*, 1993; Pais *et al.*, 2012). Similar intraspecific divergences during the Pleistocene are reported in other Iberian *Squalius* species inhabiting areas near the Gulf of Cadiz in southwestern Portugal, such as *S. aradensis* (Mesquita *et al.*, 2005). The shallow divergence time, the close mitochondrial relationship of Atlantic populations of southern Iberia with other northern basins, such as the Guadiana, and the sharing of nuclear alleles among Atlantic and Mediterranean populations of this species in southern Iberia probably as consequence of incomplete lineage sorting (Maddison, 1997; Maddison and Knowles, 2006), suggest recent colonization of the southern Iberian region by this species. DEC and S-DIVA reconstructions supported this hypothesis of colonization of both Atlantic and Mediterranean populations (Fig. 7).

Atlantic populations of *S. pyrenaicus*, as same as *S. malacitanus*, presented higher genetic diversity than did Mediterranean (Table S3 in the Supporting Information of this chapter). Some of the Atlantic populations of *S. pyrenaicus* share mitochondrial haplotypes, such is the case of Guadalquivir Basin with both Guadalete and Odiel basins, as also demonstrated the significant gene flow found among these Atlantic populations (Fig. 4 and Table S5 in the Supporting Information of this chapter). Our DEC and S-DIVA reconstructions supported two different colonization events of the Guadalete Basin coming from the Guadalquivir, being one of these dispersal events followed by a vicariance between both basins (marginal probability DEC = 0.6 and 0.54; marginal probability S-DIVA = 1.0 and 1.0; Fig. 7), which promoted the formation of an isolated mitochondrial lineage in the Guadalete Basin. Sea level change cycles during the Pleistocene affected Atlantic basins la cual produjo la formación de un linaje mitocondrial aislado en el río Guadalete. (Hernández-Molina *et al.*, 1994; Mediavilla *et al.*, 2004; Santisteba and Schulte, 2007). Nevertheless, no evidence of downstream connections between the Guadalquivir and Guadalete basins has been reported, as was borne out by our bathymetric reconstruction for the region (Fig. 2). Palaeogeographic reconstructions of the Guadalete Basin based on fluvial sediments indicate that this river has flowed through a separate incised valley since the Middle-Upper Pleistocene (Dabrio *et al.*, 2000; Mediavilla *et al.*, 2004), for this reason, other palaeogeographical events must be invoked to explain the connection between Guadalquivir and

Guadalete basins, as we discuss below. Similar events of dispersal from Guadalquivir Basin to Odiel Basin and posterior vicariance between these two basins were also estimated by DEC and S-DIVA (marginal probability DEC = 1.0; marginal probability S-DIVA = 1.0; Fig. 7).

Mediterranean populations of *S. pyrenaicus* in southern Iberia constituted two mitochondrial isolated lineages, Guadalhorce+Vélez basins, between which gene flow was high, and Guadalfeo Basin, closely related to Guadalquivir Basin (Fig. S1). However, Bayesian and ML phylogenies based on concatenated data clustered together all Mediterranean populations of this species with high support (pp=0.9; Fig. 3). Lower genetic diversity values and scarce genetic structure of Mediterranean populations of *S. pyrenaicus* are probably consequence of recent founder effects, supported by the shallow divergence times estimated (Fig. 5). DEC analysis inferred a dispersal event from Guadalquivir to Vélez (marginal probability = 0.22; Fig. 7) and a second dispersal from the ancestral area shared by Guadalquivir and Vélez to Guadalfeo, on one side, and to Guadalhorce, on the other side (marginal probability = 0.30; Fig. 7). Posterior vicariance isolated Mediterranean populations from Guadalquivir. S-DIVA analysis showed, although with low support, a colonization of *S. pyrenaicus* from Guadalquivir (Atlantic slope) to Mediterranean populations (marginal probability = 0.05; Fig. 7) and a posterior vicariance between these two areas (marginal probability = 0.17; Fig. 7).

Since no evidence of downstream connections may be invoked nor in Atlantic neither in Mediterranean populations due to the wide separation of their river mouths and the narrow continental shelf in the region (Fig. 2), colonization of southern Iberia by *S. pyrenaicus* may have taken place through headwater capture involving river basins in the Betic Mountains, a region with a complex orography and geology (Weijermans, 1991). Evidence of stream piracy during the Pleistocene is common in the eastern regions of this mountain range and in other southeastern Iberian watersheds (Wezens and Wezens, 1995; Calvache and Viseras, 1997; Pais *et al.*, 2012), and similar piracy patterns may have occurred in the Betic area involving the Guadalhorce and Vélez basins, in the Mediterranean slope, and the Guadalete and Vélez basins, in the Atlantic slope, which could explain the presence of common mitochondrial haplotypes in these basins and their tributaries (Fig. 4). Fluvial captures have also been postulated in headwaters of other southwestern Iberian basins, such as the Arade River watershed, to explain the intraspecific structure of *S. aradensis* (Mesquita *et al.*, 2005; Pais *et al.*, 2012).

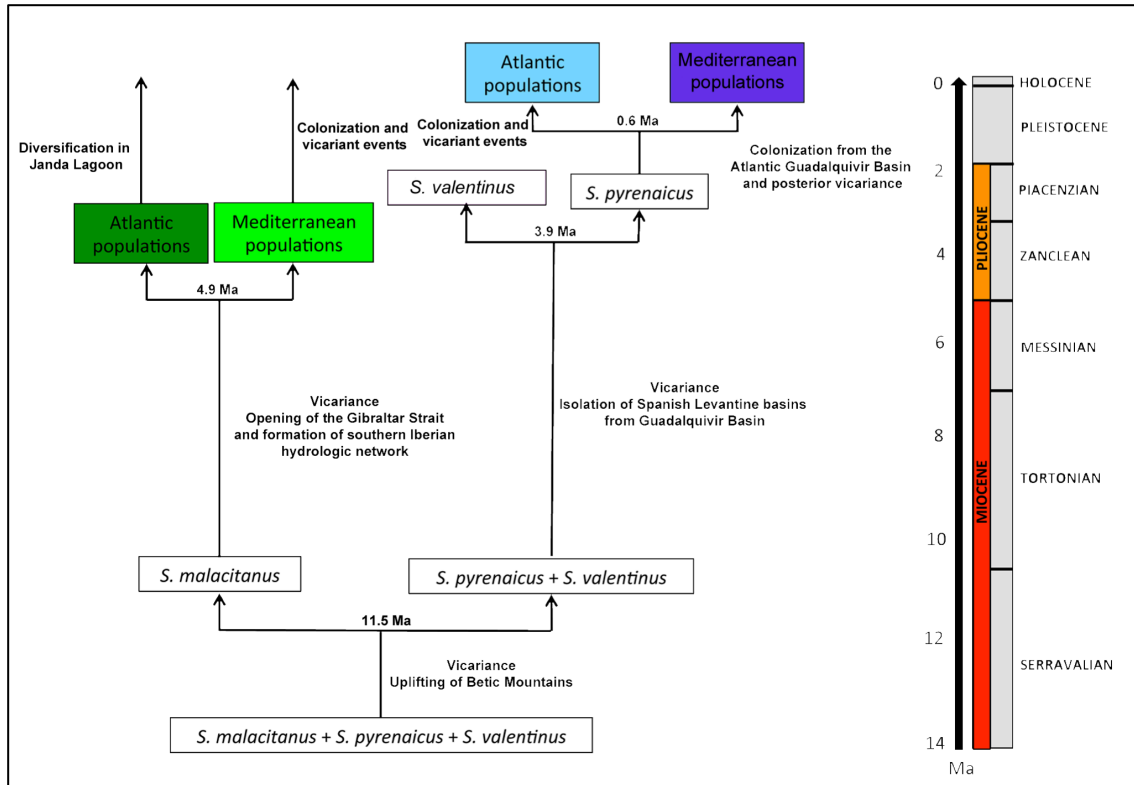


Fig. 8. Divergence and biogeographical events in *S. malacitanus* and *S. pyrenaicus*

CONCLUSIONS

Our results demonstrate the influence of Cenozoic tectonic and climatic events shaping the evolutionary history of freshwater fish fauna of the southern Iberian Peninsula (Fig. 8). We found two evolutionary patterns in two species of the cyprinid genus *Squalius* inhabiting a common area, the southern Iberian basins, mainly shaped by the time of colonization of southern Iberia. *Squalius malacitanus* appeared in the región earlier and was affected by two major tectonic processes, the uplifting of the Betic Mountains, which isolated it from other Iberian *Squalius* species, and the opening of the Gibraltar Strait, which led to the divergence of the Atlantic and Mediterranean populations.

The presence of *S. pyrenaicus* in southern Iberia is more recent than the presence of *S. malacitanus*, as a consequence of recent multiple colonization of southern Iberia from the Guadalquivir Basin through river captures since the present configuration of the hydrological network was formed, as no evidence of downstreams connections among basins was found, which is congruent with the low genetic structure found between its Atlantic and Mediterranean populations.

Our study suggests a key role of the Janda Lagoon for freshwater communities, supported by the higher genetic diversity of Atlantic populations of *S. malacitanus*, contrary to our previous expectations, as well as the lack of genetic structure in Atlantic populations of this species compared to the Mediterranean populations. The Janda Lagoon likely played an important role on genetic diversity and structure of freshwater fish species inhabiting the Iberian slope of Gibraltar Strait during the Quaternary.

AGRADECIMIENTOS

The authors thank to people that have helped in field sampling trips, among them P. Garzón, J.L. González, B. Prieto, C.P. Ornelas, C. Pedraza and J. Manzano. We also thank to L. Alcaraz for help in laboratory work and to F. Morcillo for his valuable help with bathymetric reconstruction. The Lucidus Consultancy contributed to improve the manuscript with its English edition. The Spanish Ministry of Science and Innovation has supported this study through the project CGL2010-15231BOS.

REFERENCES

- Albert EM, Zardoya R, García-Paris M. 2007. Phylogeography and speciation patterns in subterranean worm lizards of the genus *Blanus* (Amphisbaenidae: Blanidae). *Molecular Ecology*, 16(7): 1519-1531.
- Almada VC, Robalo JL, Levy A, Freyhof J, Bernardi G y Doadrio I. 2009. Phylogenetic analysis of Peri-Mediterranean blennies of the genus *Salarias*: molecular insights on the colonization of freshwaters. *Molecular Phylogenetics and Evolution*, 57(2): 924-931.
- Benito G, Thorndycraft VR, Rico M, Sánchez-Moya Y, Sopena A. 2008. Paleoflood and floodplain records from Spain: evidence for long-term climate variability and environmental changes. *Geomorphology*, 101: 68-77.
- Betzier C, Braga JC, Martín JM, Sánchez-Almazo IM, Lindhorst S. 2006. Closure of a seaway: stratigraphic record and facies (Guadix Basin, Southern Spain). *International Journal of Earth Science*, 95: 431-452.
- Böhlen J, Perdices A, Doadrio I, Economidis PS. 2006. Vicariance, colonization and fast local speciation in Asia Minor and the Balkans as revealed from the phylogeny of spined loaches (Osteichthyes, Cobitidae). *Molecular Phylogenetics and Evolution*, 39(2): 552-561.
- Braga JC, Martín JM, Quesada C. 2003. Patterns and average rates of late Neogene-Recent uplift of the Betic Cordillera, SE Spain. *Geomorphology*, 50: 3-26.
- Calvache ML, Viseras C. 1997. Long-term control mechanisms of stream piracy processes in southeast Spain. *Earth Surface Processing Landforms*, 22: 99-105.
- Calvo JP, Daams R, Morales J, López-Martínez N, Agustini J, Anadón P, Armenteros I, Cabrera L, Civis J, Corrochano A, Díaz-Molina M, Muñoz A, Pérez-García A, Pérez-González A, Portero JM, Santisteban C, Torres T, Van der Meulen AJ, Vera JA, Mein P. 1993. Up-to-date Spanish continental Neogene synthesis and paleoclimatic interpretation. *Revista de la Sociedad Geológica de España*, 6: 29-40.
- Carranza R, Harris DJ, Arnold EN, Batista V, González de la Vega JP. 2006. Phylogeography of the lacertid lizard, *Psammmodromus algerius*, in Iberia and across the Strait of Gibraltar. *Journal of Biogeography*, 33: 1279-1288.
- Carretero MA. 2008. An integrated assessment of a group with complex systematics: the Iberomaghreb lizard genus *Podarcis* (Squamata, Lacertidae). *Integrative Zoology*, 4: 247-266.

- Coelho MM, Brito RM, Pacheco TR, Figueiredo D, Pires AM. 1995. Genetic variation and divergence of *Leuciscus pyrenaicus* and *L. carolitertii* (Pisces, Cyprinidae). *Journal of Fish Biology*, 47(Suppl.A): 243-258.
- Dabrio CJ, Zazo G, Goy JL, Sierró FJ, Borja F, Lario J, González AJ, Flores JA. 2000. Depositional history of estuarine infill during the last postglacial transgression (Gulf of Cadiz, southern Spain). *Molecular Phylogenetic and Evolution*, 162: 381-404.
- Doadrio I. 1988. Phylogenetic relationships and classification of western palearctic species of the genus *Barbus* (Osteichthyes, Cyprinidae). *Aquatic Living Resources*, 3(04): 265-282
- Doadrio I, Carmona JA. 2003. Testing freshwater Lago Mare dispersal theory on the phylogeny relationships of Iberian cyprinids genera *Chondrostoma* and *Squalius* (Cypriniformes, Cyprinidae). *Graellsia*, 59(2-3): 457-473.
- Doadrio I, Carmona JA. 2004. Phylogenetic relationships and biogeography of the genus *Chondrostoma* inferred from mitochondrial DNA sequences. *Molecular Phylogenetics and Evolution*, 33: 802-81.
- Doadrio I, Carmona JA. 2006. Phylogenetic overview of the genus *Squalius* (Actinopterygii, Cyprinidae) in Iberian Peninsula, with description of two new species. *Cybum*, 30(3): 199-214.
- Doadrio I, Carmona JA, Fernández-Delgado C. 2002a. Morphometric study of the Iberian *Aphanius* (Actinopterygii, Cyprinidae), with description of a new species. *Folia Zoologica*, 51(1): 67-79.
- Doadrio I, Carmona JA, Machordom A. 2002b. Haplotype diversity and phylogenetic relationships among the Iberian barbels (*Barbus*, Cyprinidae) reveal two evolutionary lineages. *Journal of Heredity*, 93(2): 140-147.
- Doadrio I, Kottelat M, Sostoa A. 2007a. *Squalius laietanus*, a new species of cyprinid fish from north-eastern Spain and southern France (Teleostei: Cyprinidae). *Ichthyological Exploration of Freshwater*, 33: 802-815.
- Doadrio I, Perea S, Alonso F. 2007b. A new species of the genus *Squalius* Bonaparte, 1837 (Actinopterygii, Cyprinidae) from the Tagus River Basin (Central Spain). *Graellsia*, 63(1): 89-100.
- Doadrio I, Perdices A. 2005. Phylogenetic relationships among the Ibero-African cobitids (*Cobitis*, Cobitidae) based on cytochrome b sequence data. *Molecular Phylogenetics and Evolution*, 37(2): 484-493.
- Doadrio I, Perea S, Garzón-Heydt P, González JL. 2011. Ictiofauna continental española. Bases para su seguimiento. Dirección General del Medio Natural y Política Forestal. MAMRM. Madrid. 616 pp.
- Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012. A simulated annealing approach to define the genetic structure of populations. *Molecular Ecology*, 11: 2571-2581.
- Durand JD, Unlü E, Doadrio I, Pipoyan S, Templeton AR. 2000. Origin, radiation, dispersion and allopatric hybridization in the chub *Leuciscus cephalus*. *Proceedings of the Royal Society of London. Series B*. 267: 1687-1697.
- Esteban M, Braga JC, Martín JM, Santisteban C. 1996. Western Mediterranean reef complexes. In: Franseen EK, Esteban M, Ward WC, Rouchy JM (eds.). *Models for carbonate stratigraphy from Miocene reef complexes of Mediterranean regions*. SEPM. *Concepts in Sedimentology and Paleontology Series*, 5: 55-72.
- Excoffier L, Lischer HE. 2010. Arlequin suite v.3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, 10(3): 564-567.
- Filipe AFI, Araujo MB, Doadrio I, Angermeier PL, Collares-Pereira MJ. 2009. Biogeography of Iberian freshwater fishes revisited: the roles of historical versus contemporary constraints. *Journal of Biogeography*, 36(11): 2096-2110.
- Freyhof J, Kottelat M. 2008. *Squalius malacitanus*. The IUCN Red List of Threatened Species 2008: e.T135649A4170465.
- Fromhage L, Vences M, Veith M. 2004. Testing alternative vicariance scenarios in Western Mediterranean discoglossid frogs. *Molecular Phylogenetics and Evolution*, 31: 308-322.
- Gante HF, Micael J, Oliva-Paterna FJ, Doadrio I, Dowling TE, Alves MJ. 2009. Diversification within glacial refugia: tempo and mode of evolution of the polytypic fish *Barbus sclateri*. *Molecular Ecology*, 18: 3240-3255.
- García-Castellanos D, Estrada F, Jiménez-Munt I, Gorini C, Fernández M, Vergés J, De Vicente R. 2009.

- Catastrophic flood of the Mediterranean after the Messinian Salinity Crisis. *Nature*, 462(10): 778-782.
- Gómez A, Lunt DH. 2007. Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula. En: Weiss S, Ferrand N (eds.). *Phylogeography of Southern European Refugia*. Springer, Dordrecht. pp: 155-188.
- González-Peña D, Gómez-Blanco D, Reboiro-Jato M, Fernández-Riverola F, Posada D. 2010. ALTER: program-oriented conversion of DNA and protein alignments. *Nucleic Acid Research*, 38(2): 14-18.
- Goy JL, Zazo C, Dabrio CJ, Lario J, Borja F, Sierro FJ, Flores JA. 1995. *Global and Regional factors controlling changes of coastlines in southern Iberia during the Holocene. IGCP Project 367: Late Quaternary coastal records of rapid change; Application to present and future conditions*. IInd Annual Meeting, Abstracts volume, Antofagasta, Chile, November, pp: 37-38.
- Heled J, Drummond A. 2010. Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution*, 27(3): 570-580.
- Hernández-Molina FJ, Somoza L, Rey J, Pomar L. 1994. Late Pleistocene-Holocene sediments on the Spanish continental shelves: a model for very high resolution sequence stratigraphy. *Marine Geology*, 120: 129-174.
- Hudson RR, Bryant D. 2006. Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, 23(2): 254-267.
- Krijgsman W, Hilgen FJ, Raffi I, Sierro FJ, Wilson DS. 1999. Chronology, causes and progression of the Messinian Salinity Crisis. *Nature*, 400: 652-655.
- Lanfear R, Calcott B, Ho SYW, Guindon S. 2012. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution*, 29(6): 1695-1701.
- Lario J, Zazo C, Goy JL, Silva PG, Bardaji T, Cabero A, Dabrio CJ. 2011. Holocene paleotsunami catalogue of SW Iberia. *Quaternary International*, 30: 1-15.
- Librado P, Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, 25: 1451-1452.
- Loget N, Van Den Driessche J. 2006. On the origin of the Strait of Gibraltar. *Sedimentary Geology*, 188-189: 341-356.
- Lopes-Cunha M, Aboim MA, Mesquita N, Alves MJ, Doadrio I, Coelho MM. 2012. Population genetic structure in the Iberian cyprinid fish *Iberochondrostoma lemmingii* (Steindachner, 1866): disentangling species fragmentation and colonization processes. *Biological Journal of the Linnean Society*, 105: 559-572.
- Luque L, Zazo C, Goy JL, Silva PG, Bardaji T, Cabero A, Dabrio CJ. 2011. Holocene paleotsunami catalogue of SW Iberia. *Quaternary International*, 30: 1-5.
- Machordom A, Doadrio I. 2001. Evidence of Cenozoic Betic-Kabilian connection based on freshwater fish phylogeography (*Luciobarbus*, Cyprinidae). *Molecular Phylogenetics and Evolution*, 18: 252-263.
- Maddison WP. 1997. Gene tree in species trees. *Systematic Biology*, 46: 523-536.
- Maddison WP, Knowles LL. 2006. Inferring phylogeny despite incomplete lineage sorting. *Systematic Biology*, 55(1): 21-30.
- Magalhães MF. 2002. Gradientes espaciais, variabilidade ambiental e estrutura de comunidades piscícolas em rios de tipo Mediterrânico. PhD Thesis. University of Lisbon. Portugal.
- Martín JM, Braga JC, Aguirre J, Puga-Bernabéu A. 2009. History and Evolution of the North-Betic Strait (Prebetic Zone, Betic Cordillera): a narrow, early Tortonian, tidal-dominated, Atlantic-Mediterranean marine passage. *Sedimentary Geology*, 216: 80-90.
- Martín JM, Puga-Bernabéu A, Aguirre J, Braga JC. 2014. Miocene Atlantic-Mediterranean seaways in the Betic cordillera (southern Spain). *Revista de la Real Sociedad Geológica de España*, 27(1): 175-186.
- Martínez-Solano I, Gonçalves HA, Arntzen JW, García-París M. 2004. Phylogenetic relationships and biogeography of midwife toads (Discoglossidae: Alytes). *Journal of Biogeography*, 31(4): 603-618.

- Mediavilla R, Antón-López, Dabrio CJ, Perucha MA, Santisteban JL, Mediato JF, Barnolas A, Llave E. 2004. Disposición y características de los depósitos fluviales pleistocénicos del subsuelo de la bahía de Cádiz. *GeoTemas*, 6(5): 203-206.
- Mesquita N, Hänfling B, Carvalho R, Coelho MM. 2005. Phylogeography of the cyprinid *Squalius aradensis* and implications for conservation of the endemic freshwater fauna of southern Portugal. *Molecular Ecology*, 14: 1939-1954.
- Miller KG, Kominz MA, Browning JV, Wright JD, Mountain GS, Katz ME, Sugarman PJ, Cramer BS, Christie-Blick N, Pekar SF. 2005. The Phanerozoic record of global sea-level change. *Science*, 310: 1293.
- Miller MS, Allam AA, Becker TW, Di Leo JF, Wookey J. 2013. Constraints on the tectonic evolution of the westernmost Mediterranean and northwestern Africa from shear wave splitting analysis. *Earth and Planetary Science Letters*, 375(1): 234-243.
- Miraldo A, Hewitt GM, Paulo OS, Emerson BC. 2011. Phylogeography and demographic history of *Lacerta lepida* in the Iberian Peninsula: multiple refugia, range expansions and secondary contact zones. *BMC Evolutionary Biology*, 11(1): 170.
- Pais JA, Cunha PP, Pereira D, Legoinha P, Dias R, Moura D, Brum da Silveira A, Kullberg JC, González-Delgado JA. 2012. *The Paleogene and Neogene of western Iberia (Portugal). A Cenozoic record in the European Atlantic domain*. Springer Briefs in Earth Sciences. Springer. Berlin.
- Perdices A, Carmona JA, Fernández-Delgado C, Doadrio I. 2001. Nuclear and mitochondrial data reveal high genetic divergence among Atlantic and Mediterranean populations of the Iberian killifish *Aphanius iberus* (Teleostei: Cyprinodontidae). *Heredity*, 87(3): 314-324.
- Perdices A, Doadrio I, Economidisi PS, Böhlen J, Banarescu P. 2003. Pleistocene effects on the European freshwater fish fauna: double origin of the cobitid genus *Sabanejewia* in the Danube basin (Osteichthyes: Cobitidae). *Molecular Phylogenetics and Evolution*, 26(2): 289-299.
- Perdices A, Machordom A, Doadrio I. 1995. Allozyme variation of African and Iberian populations of the genus *Cobitis* Linnaeus, 1758 (Osteichthyes, Cobitidae). *Journal of Fish Biology*, 47: 707-718.
- Perea S, Böhme M, Zupancic P, Freyhof J, Sanda R, Özlüg M, Abdolí A, Doadrio I. 2010. Phylogenetic relationships and biogeographical patterns in Circum-Mediterranean subfamily Leuciscinae (Teleostei: Cyprinidae) inferred from both mitochondrial and nuclear data. *BMC Evolutionary Biology*, 10(1): 265.
- Perea S, Doadrio I. Phylogeography, historical demography and habitat suitability modeling of freshwater fishes inhabiting seasonally fluctuating Mediterranean river systems: a case study using the Iberian cyprinid *Squalius valentinus*. *Molecular Ecology*, 24: 3706-3722.
- Rambaut A, Drummond AJ. 2007. Tracer v.1.4. Available from: <http://beast.bio.ed.ac.uk/Tracer>.
- Recio JM. 2007. Medio natural y cuaternario de Gibraltar en los textos de viajero por la Península Ibérica durante los siglos XVIII y XIX. *Almoraima*, 35: 11-20.
- Ree RH, Smith SA. 2008. Maximum-Likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology*, 57: 4-14.
- Ree RH, Moore BR, Webb CO, Donoghue MJ. 2005. A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. *Evolution*, 59: 2299-2311.
- Ribera I, Vogler AP. 2004. Speciation of Iberian diving beetles in Pleistocene refugia (Coleoptera, Dysticidae). *Molecular Ecology*, 13(1): 179-193.
- Ribera I, Castro A, Díaz JA, Garrido J, Izquierdo A, Jäch MJ, Valladares LF. 2011. The geography of speciation in narrow-range endemics of the "Haenydra" lineage (Coleoptera, Hydraenidae, *Hydraena*). *Journal of Biogeography*, 38(3): 502-516.
- Rice WR. 1989. Analyzing tables of statistical tests. *Evolution*, 223-225.
- Robalo J, Almada VC, Levy A, Doadrio I. 2006. Re-examination and phylogeny of the genus *Chondrostoma* based

- on mitochondrial and nuclear data and the definition of 5 new genera. *Molecular Phylogenetics and Evolution*, 42(2): 362-376.
- Ronquist F, Telensko M, van der Mark P, Ayres DL, Darling A, Höhna, Liu L, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2.: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, 61(3): 539-542.
- Sanjur OI, Carmona JA, Doadrio I. 2003. Evolutionary and Biogeographical patterns within Iberian populations of the genus *Squalius* inferred from molecular data. *Molecular Phylogenetics and Evolution*, 29: 20-30.
- Santisteban JI, Schulte L. 2007. Fluvial networks of the Iberian Peninsula: a chronological framework. *Quaternary Science Review*, 26: 2738-2757.
- Seager R, Kiu H, Henderson N, Simpson I, Kelley C, Shaw T, Kushnir Y, Ting M. 2014. Causes of increasing aridification of the Mediterranean region in response to rising greenhouse gases. *Journal Climate*, 27: 4655-4676.
- Slatkin M, Barton NH. 1989. A comparison of three indirect methods for estimating average levels of gene flow. *Evolution*, 43(7): 1349-1368.
- Sousa V, Penha F, Collares-Pereira MJ, Chicki L, Coelho MM. 2008. Genetic structure and signature of population decrease in the critically endangered freshwater cyprinid *Chondrostoma lusitanicum*. *Conservation Genetics*, 9: 791-805.
- Sousa-Santos C, Collares-Pereira MJ, Almada V. 2007. Reading the history of a hybrid complex from its molecular record. *Molecular Phylogenetics and Evolution*, 45: 981-996.
- Sousa-Santos C, Gante HF, Robalo JI, Cunha PP, Martins A, Arruda M, Alves MJ, Almada VC. 2014. Evolutionary history and population genetics of a cyprinid fish (*Iberochondrostoma olisiponensis*) endangered by introgression from a more abundant structure. *Conservation Genetics*, 15: 665-677.
- Stamakis A. 2006. RaxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, 22(21): 2688-2690.
- Stamakis A, Blagojevic F, Nikolopoulos D, Antonopoulos C. 2007. Exploring new search algorithms and hardware for phylogenetics: RaxML meets the IBM cell. *J. VLSI Signal Process.*, 48: 271-286.
- Stephens M, Donnelly P. 2003. A comparison of Bayesian methods for haplotype reconstruction from population genotype data. *American Journal of Human Genetics*, 73: 1162-1169.
- Stephens M, Smith N, Donnelly P. 2001. A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics*, 76: 978-989.
- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA 6: molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution*, 30: 2725-2729.
- Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acid Research*, 22(22): 4673-4680.
- Weijermans R. 1991. Geology and tectonics of the Betic Zone, southeast Spain. *Earth Science Review*, 31: 153-236.
- Wenzens E, Wenzens G. 1995. The influence of Quaternary tectonics on river capture and drainages patterns in the Huerca-Overa basin, southeastern Spain. En: Lewin J, Macklin MG, Woodward JC (eds.). *Mediterranean Quaternary River Environments*. Balkema. Rotterdam., pp: 55-63.
- Yu Y, Harris AJ, He XJ. 2010. S-DIVA (statistical dispersal-vicariance analysis): a tool for inferring biogeographic histories. *Molecular Phylogenetics and Evolution*, 56(2): 848-850.
- Yu Y, Harris AJ, Blair C, He XJ. 2015. RASP (Reconstruct Ancestral State in Phylogenies): a tool for historical biogeography. *Molecular Phylogenetics and Evolution*, 56(2): 87: 46-49.
- Zardoya R, Doadrio I. 1998. Phylogenetic relationship of Iberian cyprinids: systematic and biogeographical implications. *Proceedings of the Royal Society of London, Series B*, 265: 1365-1372.

SUPPLEMENTARY MATERIAL

Table S1 . Evolutionary models selected by PartitionFinder for each analyzed gene and for the concatenated data matrix.

Gene	Total characters	Evolutionary model selected (AIC criterion)
<i>MT-CYB</i>	1140	Codon 1 st position: K80+I+G Codon 2 nd position: TVM+I+G Codon 3 rd position: GTR+G
<i>RAG1</i>	1041	Codon 1 st position: F81 Codon 2 nd position: K81uf Codon 3 rd position: TVMef
<i>MT-CYB+RAG</i> (concatenated analysis)	2181	Codon 1 st position <i>MT-CYB</i> + Codon 3 rd position <i>RAG1</i> : TVMef+I+G Codon 2 nd position <i>MT-CYB</i> : HKY Codon 3 rd position <i>MT-CYB</i> : GTR+G Codon 1 st position <i>RAG1</i> : TVM+I+G Codon 2 nd position <i>RAG1</i> : K81uf+I

Table S2. Uncorrected-*p* interpopulation genetic distances in percentage between *Squalius malacitanus* and southern Iberian populations of *Squalius pyrenaicus* for the *MT-CYB* (below diagonal) and *RAG1* (above diagonal) genes. The final two rows show intrapopulation genetic distances in percentage for *MT-CYB* and *RAG1*.

	Guadiana	Guadalquivir	Odiel	Guadalete	Celemín	Vega	Jara	Miel	Guadiaro	Guadalmina	Guadaiza	Guadalhorce	Velez	Guadalfeo
Guadiana	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Guadalquivir	0.007	-	0.003	0.003	0.006	0.007	0.006	0.007	0.005	0.006	0.006	0.003	0.003	0.006
Odiel	0.009	0.008	-	0.001	0.004	0.004	0.004	0.005	0.003	0.003	0.003	0.000	0.000	0.003
Guadalete	0.006	0.006	0.008	-	0.005	0.005	0.005	0.006	0.004	0.004	0.004	0.001	0.001	0.004
Celemín	0.072	0.073	0.074	0.071	-	0.001	0.000	0.001	0.001	0.001	0.001	0.004	0.004	0.007
Vega	0.073	0.074	0.075	0.072	0.005	-	0.001	0.002	0.002	0.002	0.002	0.005	0.005	0.007
Jara	0.073	0.074	0.075	0.072	0.004	0.002	-	0.001	0.001	0.001	0.001	0.005	0.005	0.008
Miel	0.072	0.073	0.074	0.071	0.004	0.005	0.004	-	0.002	0.002	0.002	0.003	0.003	0.006
Guadiaro	0.075	0.076	0.077	0.075	0.040	0.042	0.042	0.041	-	0.001	0.001	0.003	0.003	0.006
Guadalmina	0.078	0.079	0.079	0.078	0.042	0.044	0.044	0.042	0.005	-	0.001	0.003	0.003	0.006
Guadaiza	0.078	0.078	0.079	0.078	0.042	0.043	0.043	0.042	0.005	0.001	-	0.003	0.003	0.006
Guadalhorce	0.008	0.007	0.007	0.007	0.074	0.075	0.075	0.074	0.077	0.080	0.079	-	0.000	0.003
Velez	0.008	0.007	0.007	0.007	0.075	0.076	0.075	0.075	0.077	0.080	0.080	0.000	-	0.003
Guadalfeo	0.010	0.010	0.010	0.010	0.077	0.078	0.078	0.077	0.082	0.084	0.084	0.006	0.006	-
<i>MT-CYB</i>	0.004	0.007	0.006	0.003	0.003	0.001	0.002	0.002	0.003	0.001	0.000	0.000	0.000	0.002
<i>RAG1</i>	-	0.004	0.000	0.001	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table S3. Genetic diversity indices according to *MT-CYB* for populations of *S. malacitanus* and *S. pyrenaicus*. n = number of specimens; h = number of haplotypes; S = number of polymorphic sites; H_d = haplotype diversity; k = average number of nucleotide differences; π = nucleotide diversity; SD = standard deviation.

Population	n	h	S	H_d	H_d S.D.	k	π	π SD
<i>Squalius malacitanus</i>								
Guadiaro Basin	87	14	26	0.862	0.019	3.78	0.0033	0.00024
Guadalmina Basin	33	4	5	0.403	0.093	0.81	0.0007	0.00028
Guadaiza Basin	9	2	1	0.556	0.090	0.56	0.0005	0.00008
Barbate Basin	23	15	22	0.941	0.034	3.40	0.0030	0.00051
Jara Basin	16	7	8	0.792	0.089	1.79	0.0016	0.00029
Vega Basin	20	2	4	0.100	0.088	0.40	0.0035	0.00099
Miel Basin	12	5	8	0.556	0.017	1.56	0.0013	0.00047
<i>Squalius pyrenaicus</i>								
Guadalhorce basin	57	8	7	0.348	0.081	0.41	0.0004	0.00010
Vélez Basin	49	4	2	0.438	0.067	0.48	0.0004	0.00007
Guadalquivir Basin	125	49	66	0.936	0.013	7.87	0.0070	0.00028
Guadalete Basin	32	7	13	0.643	0.079	3.60	0.0032	0.00052
Guadalfeo	3	3	3	1.000	0.272	2.00	0.0017	0.00062
Guadiana	30	21	30	0.966	0.021	4.32	0.0038	0.00051
Odiel Basin	8	3	13	0.607	0.164	6.68	0.0058	0.00136

Table S4. Hierarchical analysis of molecular variance (AMOVA) based on the *MT-CYB* gene. df =degrees of freedom; SSD = sum of squared deviation.

Source of variation	df	SSD	Variance components	P-value	Percentage variation
<i>Squalius malacitanus</i> + <i>S. pyrenaicus</i> (One gene pool)					
Among basins	13	11494.765	25.81992	< 0.0001	93.42
Within basins	496	901.690	1.81792		6.58
FST = 0.93					
<i>S. malacitanus</i> (One gene pool)					
Among basins	6	2276.778	14.52375	< 0.0001	92.17
Within basins	199	245.650	1.23442		7.83
FST = 0.92					
<i>S. malacitanus</i> (Two gene pools: Atlantic and Mediterranean populations)					
Among groups	1	2099.879	21.20095	0.03	88.80
Among basins within groups	5	176.899	1.43881	< 0.0001	6.03
Within basins	199	245.650	1.23442	< 0.0001	5.17
FST = 0.95					
FSC = 0.54					
FCT = 0.89					
<i>S. pyrenaicus</i> (One gene pool)					
Among basins	6	390.936	1.66053	< 0.0001	42.91
Within basins	297	656.041	2.20889		57.09
FST = 0.43					
<i>S. pyrenaicus</i> (Two gene pools: Atlantic and Mediterranean populations)					
Among groups	1	262.552	1.52103	0.028	34.14
Among basins within groups	5	128.384	0.72540	< 0.0001	16.28
Within basins	297	656.041	2.20889	< 0.0001	49.58
FST = 0.50					
FSC = 0.25					
FCT = 0.34					

Table S5. Pairwise Φ_{ST} values (below diagonal) and gene flow (above diagonal) for the *MT-CYB* gene between populations of *S. malacitanus* and southern Iberian populations of *S. pyrenaicus*. All pairwise Φ_{ST} comparisons were significant except for that comparing Guadalmina and Guadaiza basins. Gene flow is estimated as number of migrants contributing alleles each generation (Nm). $Nm > 1$ in bold.

	Guadiana	Guadalquivir	Odiel	Guadalete	Celemin	Vega	Jara	Miel	Guadiaro	Guadalmina	Guadaiza	Guadalhorce	Velez	Guadalfeo
Guadiana	-	1.74	0.56	0.83	0.02	0.02	0.02	0.02	0.02	0.01	0.02	0.12	0.14	0.25
Guadalquivir	0.22	-	1.71	2.67	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.82	0.85	0.68
Odiel	0.47	0.22	-	0.60	0.02	0.01	0.02	0.02	0.02	0.01	0.02	0.14	0.15	0.42
Guadalete	0.38	0.16	0.45	-	0.02	0.01	0.02	0.02	0.02	0.01	0.02	0.14	0.15	0.20
Celemin	0.95	0.92	0.95	0.96	-	0.29	0.55	0.75	0.03	0.02	0.02	0.02	0.01	0.02
Vega	0.97	0.92	0.97	0.97	0.63	-	0.52	0.10	0.03	0.01	0.00	0.00	0.00	0.00
Jara	0.96	0.92	0.96	0.96	0.47	0.48	-	0.26	0.03	0.01	0.01	0.00	0.00	0.01
Miel	0.96	0.92	0.96	0.96	0.40	0.82	0.65	-	0.04	0.01	0.01	0.00	0.00	0.01
Guadiaro	0.95	0.93	0.95	0.96	0.92	0.93	0.93	0.93	-	0.43	0.62	0.01	0.01	0.02
Guadalmina	0.97	0.93	0.98	0.97	0.96	0.99	0.98	0.98	0.53	-	4.84	0.00	0.00	0.00
Guadaiza	0.96	0.92	0.96	0.97	0.94	0.99	0.97	0.97	0.45	0.09	-	0.02	0.00	0.05
Guadalhorce	0.82	0.38	0.78	0.78	0.98	0.99	0.99	0.99	0.97	0.99	0.99	-	3.29	0.04
Velez	0.80	0.37	0.77	0.77	0.98	0.99	0.99	0.99	0.97	0.99	0.99	0.13	-	0.04
Guadalfeo	0.67	0.42	0.54	0.71	0.96	0.99	0.98	0.97	0.96	0.99	0.99	0.91	0.91	-

Fig. S1. Phylogenetic trees of independent *MT-CYB* gene on BI and ML. Posterior probability values for BI and bootstrap values for ML (1000 replicates) are shown before and after slash, respectively. Labels indicate rivers systems.

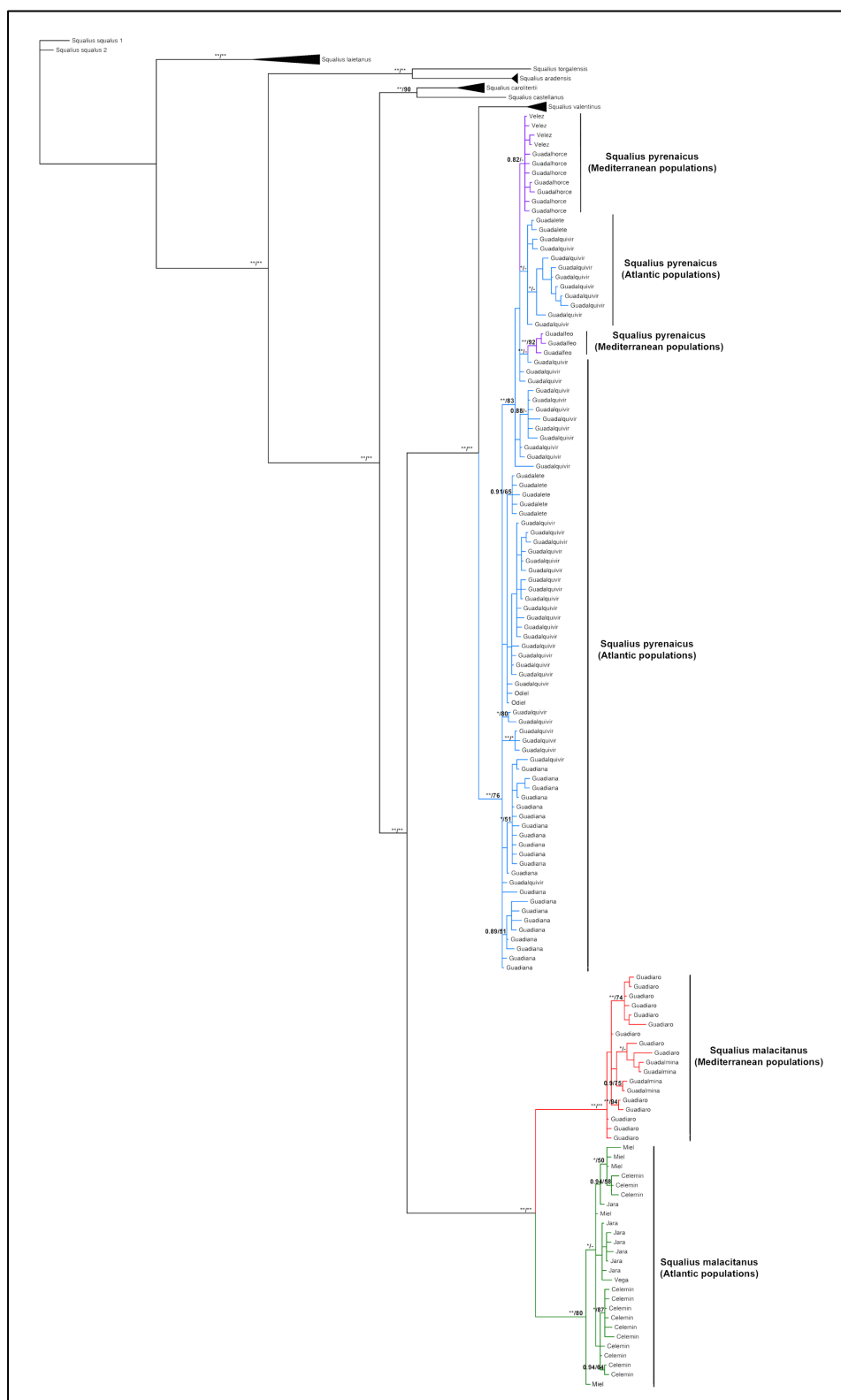
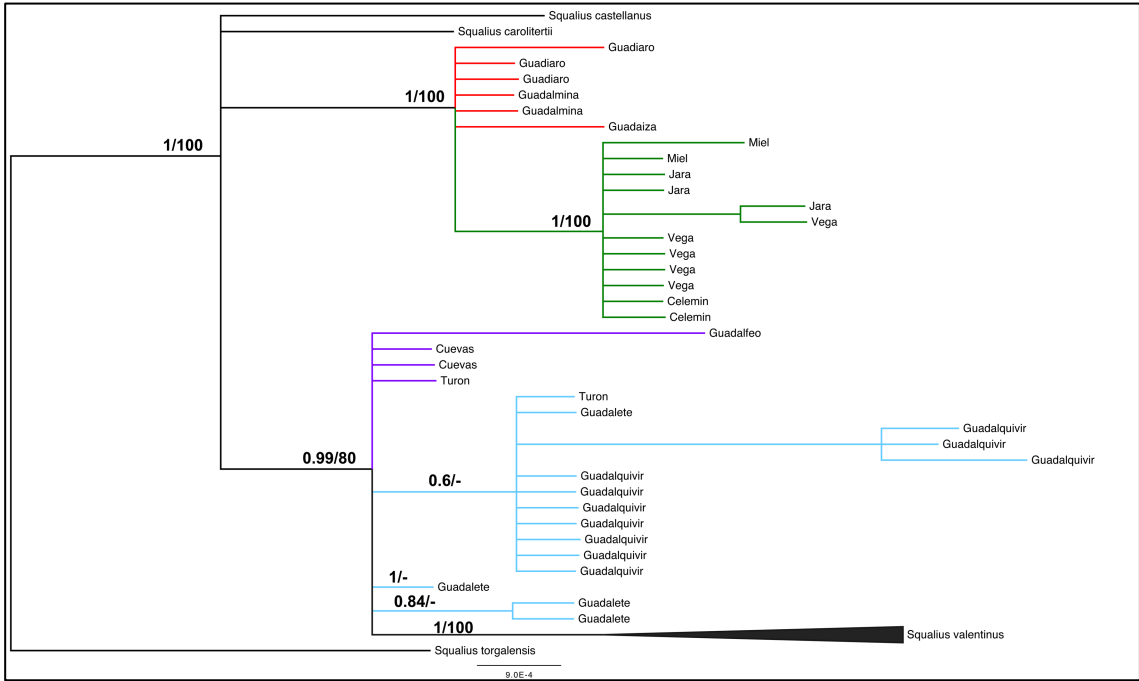


Fig. S2. Phylogenetic trees of independent *RAG1* gene on BI and ML. Posterior probability values for BI and bootstrap values for ML (1000 replicates) are shown before and after slash, respectively. Labels indicate rivers systems.



CAPÍTULO II

Influence of Iberian paleogeography on the evolutionary history of two Iberian endemic *Squalius* species (Actinopterygii, Cyprinidae): a multilocus phylogenetic and biogeographical approach.

(Manuscrito en preparación)

RESUMEN

La península ibérica es una de las regiones geológicas más antiguas en Europa, habiéndose formado como resultado del cierre del Océano de Tethys a consecuencia de colisión entre varias placas tectónicas. Además, presenta una orografía muy compleja producto de varios eventos geológicos relacionados con dos episodios orogénicos ocurridos desde el Devoniano Tardío (orogenia Varisca) hasta el Cenozoico (orogenia Alpina). Dentro de los peces ibéricos, el género *Squalius* es uno de los más diversificados, perteneciendo nueve de las diez especies endémicas de la península al linaje Mediterráneo. A pesar de que este género ha sido estudiado desde hace al menos dos décadas, varias cuestiones evolutivas siguen pendientes por resolver. De esta forma, en este estudio se llevó a cabo un análisis filogenético y filogeográfico de las dos especies con el área de distribución más amplia en la península, *S. carolitertii* y *S. pyrenaicus*, con el fin de descifrar las relaciones entre los distintos linajes de estas especies, así como patrones de estructura y diversidad genética. Además, se realizaron análisis de biogeografía histórica para desentrañar los principales patrones biogeográficos seguidos por las poblaciones de estas dos especies. Resultados de este estudio revelaron una subestimación de la diversidad real del género en la península ibérica al identificar linajes evolutivos divergentes dentro de estas dos especies, hecho que pone de manifiesto la necesidad de realizar futuros trabajos taxonómicos de estas especies ibéricas. También se identificaron diferentes patrones de estructura y diversidad genética entre *S. carolitertii* y *S. pyrenaicus*, asociados a diferentes factores históricos como las glaciaciones pleistocénicas o la estacionalidad mediterránea respectivamente. De acuerdo a los análisis biogeográficos realizados, se estimaron varios procesos de vicarianza y dispersión ocurridos en los diferentes linajes de *S. carolitertii* y *S. pyrenaicus*, ligados a la configuración paleogeográfica de la península ibérica durante el Cenozoico, especialmente importante en peces de agua dulce primarios como es el caso del género *Squalius*.

Palabras clave: árboles de especies, biogeografía histórica, discordancia mitonuclear, filogenia, *S. pyrenaicus*, *S. carolitertii*.

ABSTRACT

The Iberian Peninsula is one of the most ancient geological regions in Europe, having formed as a result of the Tertiary closure of the Tethys Ocean due to tectonic plates collision. Besides, it has a complex orography as consequence of several geological episodes related to two orogenic episodes occurred since Late Devonian (Variscan orogeny) to the Cenozoic (Alpine orogeny). Within the Iberian freshwater fauna, the genus *Squalius* is one of the most diversified, belonging nine out of ten endemic Iberian species to the Mediterranean lineage. In spite of this genus has been studied since the last two decades, several evolutionary questions remain open. Therefore, in this study we carried out phylogenetic and phylogeographical analyses for the two species with the widest distribution range in the peninsula, *S. carolitertii* and *S. pyrenaicus*, in order to decipher relationships among lineages within these species as well as patterns of genetic structure and diversity. Furthermore, historical biogeographic analyses were performed with the aim to unravel the main biogeographic patterns followed by the populations of these two species. Outcomes from this study revealed an underestimation of the real diversity of the genus *Squalius* in the Iberian Peninsula, identifying divergent evolutionary lineages within these two species supported by mitochondrial and nuclear data, which highlights the need for further taxonomic studies within Iberian *Squalius* species. We also identified different patterns of genetic structure and diversity between *S. carolitertii* and *S. pyrenaicus* associated to different historical factors such as Pleistocene glaciations and Mediterranean seasonality respectively. According to biogeographic analyses accomplished, several vicariance and dispersal processes were estimated to have occurred among *S. carolitertii* and the different lineages of *S. pyrenaicus* linked to the paleogeographical configuration of the Iberian Peninsula during the Cenozoic period, especially important in primary freshwater fishes as the genus *Squalius*.

Keywords: mitonuclear discordance, historical biogeography, phylogeny, species tree analyses, *S. pyrenaicus*, *S. carolitertii*.

INTRODUCTION

The Iberian Peninsula constitutes a particular and ancient geological unit in Europe having formed as a result of the Tertiary closure of the Tethys Ocean during the collision of India, Arabia and Africa with Eurasia (Dercourt *et al.*, 1986). The western part of the Iberian Peninsula is the oldest and it was arisen during the Variscan orogeny, which was originated by the collision between Laurasia and Gondwana and lasted from Late Devonian (ca. 370 Ma) to Carboniferous (ca. 290) periods (Torné *et al.*, 2015). The eastern and southern Iberian regions are younger; these areas had a tectonic origin during the Cenozoic as consequence of the Alpine collision that started at the Late Cretaceous (ca. 33.8 Ma) and continue until recent times giving rise to the main eastern and southern Iberian Mountain chains (Pyrenees, Iberian Chain and Betic Mountains) (Torné *et al.*, 2015). Therefore, the tectonic evolution of the Iberian microplate has led to drastic changes of its topography since Cenozoic to the present day related to a lithospheric folding (Andeweg, 2002; Vergés and Fernández, 2006; Cloetingh and Ziegler, 2007; De Vicente *et al.*, 2011). Within the Cenozoic period, tectonic activity was the highest during the Oligocene and decreased during the Early Miocene. During the Late Miocene and Pliocene some structures underwent tectonic rejuvenation (De Bruijne and Andriessen, 2002). The westernmost areas of the peninsula still show activity in the present period (Cabral and Ribeiro, 1990; Ferreira-Soares *et al.*, 2005).

Associated to these complex tectonic periods driven by the two orogenic episodes (Variscan and Alpine) is the history of the hydrological network of the Iberian Peninsula. Therefore, after the Variscan orogeny, a Mesozoic dismembering of the majority of Iberian terrains resulted in the opening of numerous extensional basins with different orientations that were later reconfigured after a significant intraplate deformation occurred during the Alpine orogeny and culminated with the origin of large endorheic basins in the Miocene that at the end acquire freshwater lacustrine conditions (Pinheiro *et al.*, 1996; De Vicente *et al.*, 2011; Torné *et al.*, 2015). These complex tectonic movements had also a strong influence into the evolutionary history of the Iberian freshwater communities, among them fish fauna, which in fact is very ancient (Doadrio *et al.*, 2011). The development of Cenozoic endorheic sedimentary basins in the Iberian Peninsula was very relevant for the Iberian freshwater fish fauna. These Miocene endorheic basins have given rise to large evolutionary lineages through vicariance in the majority of Iberian freshwater fish genera (Carmona *et al.*, 2000; Doadrio and Carmona, 2003; Sanjur *et al.*, 2003; Robalo *et al.*, 2006; Filipe *et al.*, 2009). The endorheic lagoons isolated freshwater fish fauna promoting speciation processes that continued in time until the final configuration of the current Iberian hydrological basins in the Plio-Pleistocene periods when a change from endorheic to exorheic conditions took place (Calvo *et al.*, 1993; De Vicente *et al.*, 2011; Cunha *et al.*, 2012;

Pais *et al.*, 2012), in which those speciation processes were completed and intraspecific diversification within the different freshwater fish species begun (Zardoya and Doadrio, 1998; Robalo *et al.*, 2006; Filipe *et al.*, 2009; Lopes-Cunha *et al.*, 2012; Perea *et al.*, 2016).

Within the genera of Iberian freshwater fish, the genus *Squalius* is one of the most ancient, fossils of this genus from the Miocene period are known from Portugal (*Leuciscus antunesi* Gaudant, 1977; Cabrera and Gaudant, 1985; de la Peña, 1995) and some other fossil remains from Oligocene could be ascribed to this genus on the basis of its close similarity with *S. alburnoides* morphology (*Rutilus antiquus*; Cabrera and Gaudant, 1985). The genus *Squalius* is also one of the most diversified genera in the peninsula, inhabiting in almost all the main Iberian basins (Almada and Sousa-Santos, 2010; Doadrio *et al.*, 2011; Sousa-Santos *et al.*, 2016). The diversification of the species belonging to this genus in the Iberian Peninsula is closely related to the Cenozoic endorheic basins abovementioned. The two largest basins were the Tagus and the Douro and two of the three main vicariant Iberian lineages of *Squalius* belonging to the Mediterranean group *sensu* Sanjur *et al.* (2003) have been originated in these areas, with small overlapping of their distribution ranges probably as consequence of recent secondary contacts. One of these lineages is that constituted by *S. pyrenaicus* / *S. valentinus* / *S. malacitanus*, three species that in the present day occur respectively in the Tagus Basin, the Spanish Levantine area and the southern Spanish region. The second lineage is that composed of *S. carolitertii* and *S. castellanus*, nowadays inhabiting in the Douro and other northwestern Iberian basins the former species and in the Upper Tagus Basin the latter one, respectively. The third lineage includes the species *S. aradensis* and *S. torgalensis*, which are currently distributed in the southwestern small Portuguese basins (Almada and Sousa-Santos, 2010; Perea *et al.*, 2016; Sousa-Santos *et al.*, 2016).

Out of a total of the nine species of the genus *Squalius* occurring in the Iberian Peninsula, taking into account *S. laietanus* from the Ebro Basin, belonging to the Euroasiatic group *sensu* Sanjur *et al.* (2003), and the species of hybridogenetic origin *S. alburnoides* (Cunha *et al.*, 2004), both not considered in this study, *S. pyrenaicus* is the one with the wider distribution range. This species occupies mainly the southern half of the Iberian Peninsula and is present in Tagus, Guadiana, Guadalquivir, Segura, Odiel, Sado, Gilão, as well as in small Portuguese basins close to the mouth of the Tagus, small southern Spanish basins close to the Gibraltar Strait and punctual areas of the Ebro Basin (Doadrio *et al.*, 2011; Sousa-Santos *et al.*, 2016; Fig. 1). The origin of *S. pyrenaicus* is related to the Cenozoic endorheic basins, when the lineage of *S. carolitertii* / *S. castellanus* was isolated from the lineage of *S. pyrenaicus* / *S. valentinus* / *S. malacitanus* in the Middle Miocene period (Perea *et al.*, 2016). Later on, the end of the uplifting of the Betic Mountains and the formation of the southern Iberian hydrological network in the

Serravalian and Tortonian periods separated *S. pyrenaicus* / *S. valentinus* from *S. malacitanus* (Martin *et al.*, 2009, 2014; Perea *et al.*, 2016) and, posteriorly, the separation of the Tagus and Jucar basins and the formation of the current Spanish Levantine drainages associated to the opening of the Valencia Trough resulted in the divergence of *S. pyrenaicus* from *S. valentinus* in the Pliocene period (Janssen *et al.*, 1993; Doadrio and Carmona, 2006; Perea and Doadrio, 2015; Perea *et al.*, 2016). The species *Squalius carolitertii* has the second broader distribution range in the Iberian Peninsula, inhabiting mainly in Mondego, Douro and Miño basins as well as in other small basins from the northwestern quadrant of the Iberian Peninsula (Doadrio *et al.*, 2011; Sousa-Santos *et al.*, 2016; Fig. 2). Such in the case of *S. pyrenaicus*, the origin of the species *Squalius carolitertii* is also associated to the Cenozoic endorheic basins and later isolation from the species *S. castellanus* in the Middle Miocene (Doadrio *et al.*, 2007b; Perea *et al.*, 2016).

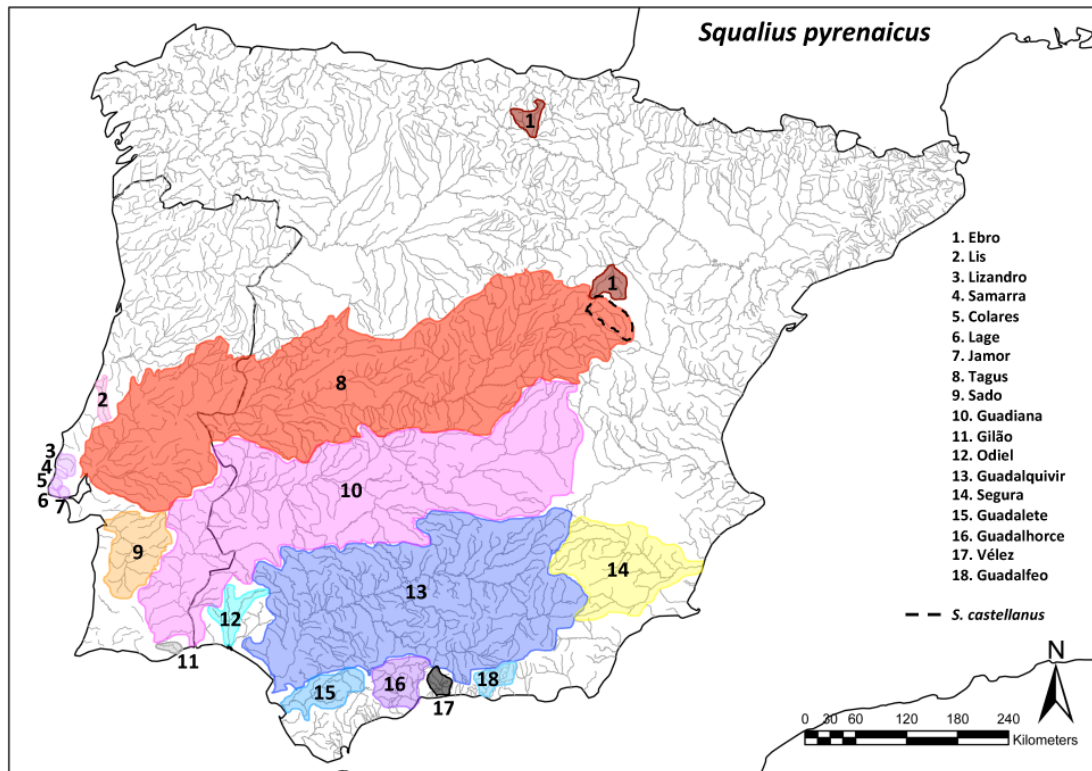


Fig. 1. Distribution range of *Squalius pyrenaicus*.

Whereas the speciation patterns within the genus *Squalius* in the Iberian Peninsula seems to be well established (Sanjur *et al.*, 2003; Doadrio and Carmona, 2003, 2004, 2006; Doadrio *et al.*, 2007a,b, Almada and Sousa-Santos, 2010; Perea *et al.*, 2016), it is not clear the intraspecific diversification of all Iberian species of this genus. Until now, it has been carried out only some intraspecific studies for Iberian *Squalius* species but these approaches do not cover their complete distribution ranges (Mesquita *et al.*, 2005, 2007; Almada and Sousa-Santos, 2010;

Henriques *et al.*, 2010; Perea *et al.*, 2016). Some previous phylogenetic studies demonstrated a high mitochondrial structuration in the populations of *S. pyrenaicus* related to hydrological basins (Doadrio and Carmona, 2006; Doadrio *et al.*, 2007b), but these studies only included a few individuals and no nuclear genes have been explored. A more exhaustive study of the genus *Squalius* in the southern Iberian basins, recognized as an independent ichthyologic region (Doadrio, 1988; Filipe *et al.*, 2009), using the *MT-CYB* gene and the nuclear *RAG1* gene also showed a high structuration between Atlantic and Mediterranean populations of *S. pyrenaicus* in this area (Perea *et al.*, 2016). A high intraspecific mitochondrial structuration within *S. carolitertii* has also been observed (Almada and Sousa-Santos, 2010). But a comprehensive mitochondrial and nuclear study of these two species in their entire distribution range has not been performed yet.

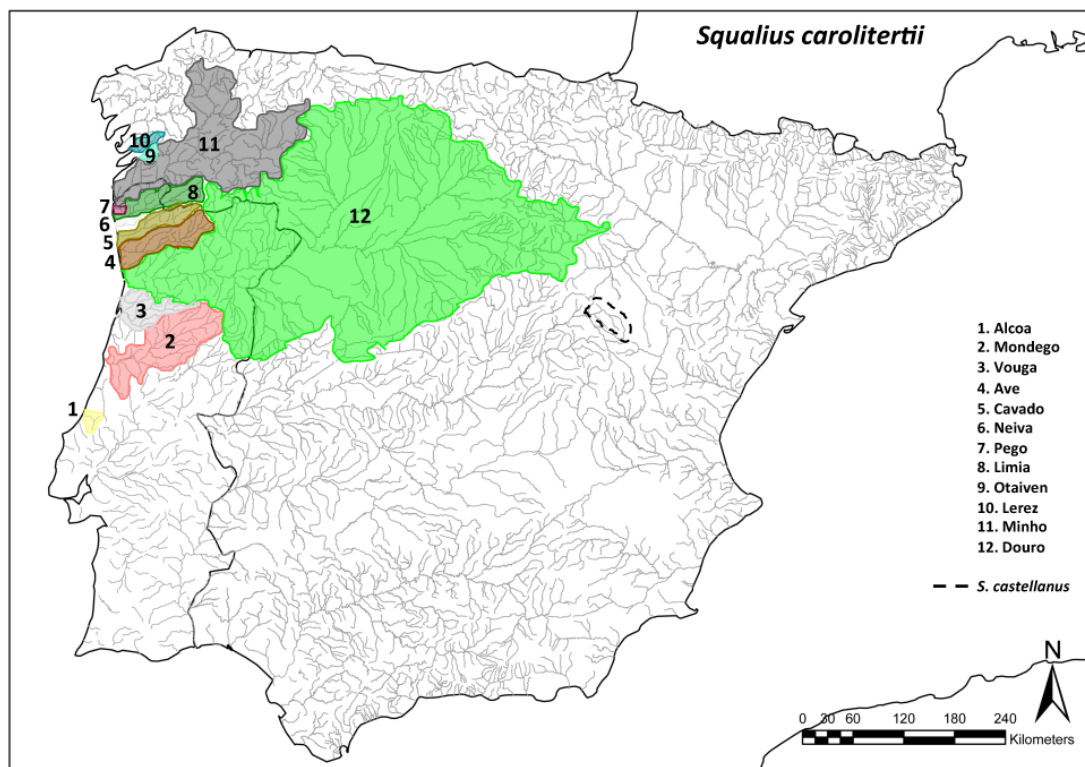


Fig. 2. Distribution range of *Squalius carolitertii*.

In the present study we carry out a phylogenetic and phylogeographical analysis of the populations of *S. pyrenaicus* and *S. carolitertii* on the basis of a broad geographical sampling covering the entire distribution range of these species in the Iberian Peninsula. We hypothesize that the evolutionary history of these two species, primary freshwater fishes (Myers, 1938), is conditioned by the formation of the Iberian hydrological network and, hence, the phylogenetic and phylogeographic relationships of these species will reflect these paleogeographical events and a high geographical structuration. We also hypothesize a higher genetic diversity in the largest

basins Douro, Tagus, Guadiana and Guadalquivir as consequence of historical factors associated to the Miocene sedimentary basins leading into ancient lineages of these Iberian endemic species. We assess these hypotheses here by analysing fragments of one mitochondrial gene (*MT-CYB*) and six nuclear genes (see Material and Methods section), with the aim to reconstruct the phylogenetic relationships of the populations of *S. pyrenaicus* and *S. carolitertii* and the biogeographic history followed by these species through ancestral area reconstruction approaches.

Therefore, the aims of this study were as follow: i) to infer the phylogenetic relationships of *S. pyrenaicus* and *S. carolitertii* populations; ii) to assess geographical congruence in the genetic structure of their populations; iii) to estimate the main genetic diversity parameters in both species with the aim to determine specific patterns in genetic diversity; iv) to estimate divergence times among *S. pyrenaicus* populations; and v) to reconstruct the biogeographic history of these species in the Iberian Peninsula.

MATERIAL AND METHODS

1. Sampling and Laboratory procedures

A total of 935 specimens of *Squalius pyrenaicus* and 718 of *S. carolitertii* from its complete distribution range were analyzed (147 localities belonging to 30 drainage basins; Table S1 in Supplementary material; Figs. 1 and 2). Additional individuals of the different Iberian *Squalius* species except for *S. alburnoides* and *S. laietanus* were included in the analyses in order to resolve the phylogenetic position of populations of *S. pyrenaicus* and *S. carolitertii* (Table 1). *Squalius aradensis* and *S. torgalensis* were used as outgroup for rooting the phylogenetic tree based on previous studies (Almada and Sousa-Santos, 2010; Perea *et al.*, 2016). Fish were captured by electrofishing under local authority permission, fin clipped, and returned to the stream after morphological identification. Fin clips were preserved in 95% ethanol. All voucher samples were stored at the DNA collection of the Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain as well as at the ISPA-MARE collection.

A fragment of the mitochondrial DNA cytochrome *b* gene (*MT-CYB*; 720 bp) of 1656 individuals and a fragment of the third exon of the nuclear Recombination Activating Gene 1 (*RAG1*; 832 bp) of 456 individuals (912 phased alleles), representing all *S. pyrenaicus* and *S. carolitertii* populations (Table S1), were amplified in order to assess the genetic structure and diversity *S. pyrenaicus* populations. Five additional nuclear genes were amplified for at least two

individuals representative of all populations with the aim to obtain a well-resolved and accurate phylogenetic tree (*S7*, *EFA1α*, *EGR2b*, *RHO*, *ACTB*). For constructing this phylogenetic tree, the final alignment length including all mitochondrial and nuclear data for 105 individuals contained 5586 bp (Table 1). PCR primers and references are cited in Table 1. PCR protocols followed conditions explained in each bibliographic reference from which primers were obtained. After checking PCR products on 1% agarose gels, they were purified by ExoSAP-IT (USB, Cleveland, OH, USA) and directly sequenced on MACROGEN Inc. (Amsterdam, The Netherlands; <http://www.macrogen.com>). All new sequences (haplotypes and alleles) were deposited in the GenBank database (Accession Numbers: xxx-xxx).

Table 1. Gene performance.

Gene	Length (bp)	Primers used in PCR (Reference)
<i>MT-CYB</i>	720	Glu F / Glu R (Machordom and Doadrio, 2001)
<i>RAG1</i>	832	RAG1F / RAG9R (Perdices <i>et al.</i> , 2005)
<i>S7</i>	907	S71F / S72R (Chow and Hazama, 1998)
<i>EFA1α</i>	655	Ef1af2 / Ef1aR2 (Waap <i>et al.</i> , 2010)
<i>RHO</i>	792	RhForward / RhReverse (Waap <i>et al.</i> , 2010)
<i>ACTB</i>	933	B-actin-Fwd / B-actin-Rev (Robalo <i>et al.</i> , 2006)
<i>EGR2b</i>	747	Egr2bF / Egr2bR (Waap <i>et al.</i> , 2010)

2. Phylogeny and phylogeography

Sequences were aligned using the default pairwise and multiple alignment parameters in Clustal W (Thompson *et al.*, 1994) implemented in MEGA software v.7.0 (Tamura *et al.*, 2013). Nuclear data was phased with the PHASE algorithm (Stephens *et al.*, 2001; Stephens and Donnelly, 2003) in DNAsp v.5.0 (Librado and Rozas, 2009) using a probability threshold of 0.9 to resolve alleles. Phases of heterozygous individuals by indels were reconstructed using the method of Sousa-Santos *et al.*, (2005). Recombination of the nuclear *RAG1* gene, used for inferring genetic structure of populations, was tested using the PHI test in Splitstree v. 4.13 (Huson and Bryant, 2006) and no evidence of recombination was found ($p > 0.05$).

Mitochondrial phylogeny was based on a subset of 105 individuals of *S. carolitertii* and *S. pyrenaicus* from all basins. Two concatenated phylogenies were performed for those 105 individuals on the basis of datasets including specimens amplified for all nuclear genes and for mitochondrial plus all nuclear genes. Homogeneity partition test (ILD) for nuclear genes was carried out in PAUP* (Swofford *et al.*, 2003). For MT-CYB analysis, best-fit model of molecular evolutions was obtained with jModeltest v.2.0 (Darriba *et al.*, 2014). For concatenated analyses,

the best-fit model of molecular evolution for each gene and codon position was estimated via AIC in PartitionFinder v. 1.1.1 (Lanfear *et al.*, 2012). Best partition schemes used in different phylogenetic analyses based on the selected evolutionary models are presented in Table 2. Bayesian inference (BI) was performed using MrBayes v.3.2 (Ronquist *et al.*, 2012). Two simultaneous analyses were run for 50⁶ generations, each with four MCMC chains sampling every 100 generations. Convergence was checked with Tracer v.1.6. (Rambaut *et al.*, 2014). After discarding the first 10% of generations as burn-in, the 50% majority rule consensus tree and posterior probabilities were obtained. Phylogenetic analyses of independent *MT-CYB* and concatenated datasets were also conducted using Maximum Likelihood (ML) using the software RaxML implemented in the Trex-online server (Stamakis, 2006) using the substitution model schemes represented in Table 2 and the rapid bootstrap algorithm (Stamakis *et al.*, 2007). Node confidence in ML analyses was estimated by rapid bootstrapping using a random seed (1000 replicates). Uncorrected *p*-distances for *MT-CYB* were estimated in MEGA v.7.0 (Tamura *et al.*, 2013) in order to quantify genetic differences among species and among populations within species. To assess the phylogeographic history of *Squalius carolitertii* and *S. pyrenaicus* in their entire distribution range we reconstructed two independent networks for *MT-CYB* and *RAG1* sequences using the software PopART (<http://popart.otago.ac.nz>). Networks were inferred using the TCS method (Clement *et al.*, 2002).

Table 2. Evolutionary model partition schemes used for different phylogenetic performance.

Phylogenetic analysis	Best partition scheme
<i>MT-CYB</i>	TIM2 + I + G
Concatenated <i>mt</i> + nuclear	7 partitions (MT-CYB : GTR+G; RAG1 : TVMef+G; S7 : TrN+I+G; EFA1 : TVM+G; RHO : TVM+G; ACTB : K81uf+G; EGR2b : GTR+G)
Concatenated nuclear	5 partitions (RAG1 : TVMef+G; S7 : TrN+I+G; EFA1 + ACTB : TVM+G; RHO : TVM+G; EGR2b : GTR+G)

3. Genetic diversity and genetic structure

We estimated levels of genetic diversity based on *MT-CYB*, including number of haplotypes (H), haplotype diversity (H_d) and nucleotide diversity (π), using DnaSP v. 5.0 (Librado and Rozas, 2009) for *Squalius pyrenaicus* and *S. carolitertii* populations independently. To assess the relative contribution of genetic variation to structure within and between populations in *S. carolitertii* and *S. pyrenaicus*, we performed an analysis of molecular variance (AMOVA) implemented in Arlequin v.3.5. (Excoffier and Lischer, 2010) and tested for genetic structure among hydrological

basins as well as among evolutionary lineages identified in each species. The genetic differentiation among populations was addressed through *MT-CYB* Φ_{ST} pairwise comparisons (Hudson *et al.*, 1992) and estimates of gene flow as virtual number of migrants (*Nm*) exchanged among populations per generation (Slatkin and Barton, 1989), using Arlequin v. 3.5. (Excoffier and Lischer, 2010). Significant deviations from the null hypothesis of no differentiation were assessed with 10 000 permutation tests. As multiple paired tests were performed, *p*-values were adjusted by Bonferroni's correction (Rice, 1989).

4. Species tree and divergence time estimation

Two multilocus species tree analyses were carried out using *BEAST implemented in Beast software v.1.8.3 (Heled and Drummond, 2010; Drummond *et al.*, 2012) using the concatenated mitochondrial + nuclear data matrix and the only-nuclear concatenated data matrix for all analyzed populations of *Squalius carolitertii* and *S. pyrenaicus*. This analysis incorporates the multispecies coalescent model into the Bayesian framework of species delimitation in order to deal with the stochasticity of the process of lineage sorting, demography and divergence time estimation providing the species tree as well as the underlying gene trees at the same time (Heled and Drummond, 2010). Therefore, we tested species boundaries between formally described Iberian *Squalius* species along with differentiated and highly supported genetic lineages within *S. carolitertii* and *S. pyrenaicus*.

Species tree analyses were also used to estimate divergence time among Iberian *Squalius* species as well as among *S. carolitertii* and *S. pyrenaicus* populations using a lognormal relaxed molecular clock and a Yule speciation model. The molecular clock was calibrated incorporating to the species tree the fossil species *Leuciscus antunesi* (13-5-14.5 Ma; Cabrera y Gaudant, 1985) from a deposit located in the Tagus Basin in Portugal. We placed this fossil in the phylogenetic tree as common ancestor for all *Squalius* populations from the Tagus basin. A lognormal prior was set for fossil calibration. Besides, a normal prior on substitution rate was applied to the *MT-CYB* gene tree in the concatenated mitochondrial + nuclear matrix, using the evolutionary rate of 0.4% per lineage per million years based on a previous study on the subfamily Leuciscinae (Perea *et al.*, 2010). MCMC analyses were run for 100⁶ generations, with parameters logged every 10000 generations. The remaining parameters were used as default. Outputs from *BEAST analyses were examined in Tracer 1.6, and the results were summarized using TreeAnnotator 1.8.3 (Drummond *et al.*, 2012). A cloudogram of the species trees was created from the posterior distribution of species tree with DensiTree (Bouckaert, 2010).

5. Ancestral area reconstruction

We reconstructed the ancestral states of the distribution ranges of the populations of the species *Squalius carolitertii* and *S. pyrenaicus* using the BayArea method (Bayesian Inference for discrete areas; Landis *et al.* 2013) implemented in RASP v.3.1 software (Yu *et al.*, 2015), as this method allows dealing with a high number of target areas. The species *S. castellanus*, *S. malacitanus* and *S. valentinus* were included in the analysis due to their phylogenetic position relative to *S. carolitertii* and *S. pyrenaicus*. The ultrametric and dichotomous tree obtained for the *MT-CYB* gene tree in the *BEAST analysis of concatenated + nuclear matrix, containing all analyzed populations, was used as the tree topology on which mapping ancestral areas. Tips of the ultrametric tree were code by areas (basins) in order to reconstruct ancestral states. Area selection for biogeographic reconstruction was based on previous Iberian biogeographic regionalization on the basis of freshwater fish fauna (see Filipe *et al.*, 2009). Therefore, twenty-two areas were selected as follow: A) *Squalius malacitanus*, B) *Squalius valentinus*, C) Guadalquivir, D) Guadalete, E) Odiel, F) Segura, G) Guadalfeo, H) Guadalhorce-Velez, I) Guadiana-Gilão basins, J) small basins close to the Tagus mouth, K) Tagus Basin, L) Ebro, M) Lis, N) Sado, O) Douro Basin, P) Pego, Q) Lézrez-Otaiven, R) Minho, S) Vouga, T) Neiva, U) Limia, V) Mondego.

RESULTS

1. Gene trees vs species trees

Mitochondrial phylogenetic tree based on *MT-CYB* gene rendered congruent topologies in both BI and ML analyses (Fig. 3). Although only a fragment of *MT-CYB* gene was analysed (720 bp), all Iberian *Squalius* species were recovered as monophyletic with high support with the exception of *S. carolitertii*, in which statistical support was moderate, probably because the short length of the analysed fragment (pp=0.8; bootstrap=45) and the species *S. pyrenaicus*, which showed three different mitochondrial lineages: one highly supported group (pp=0.96; bootstrap=78) constituted by Tagus Basin and small basins close to its mouth together with Ebro Basin; a second moderately supported group (pp=0.88; bootstrap=66) that clustered populations from Guadiana, Gilão, Guadalquivir, Odiel, Guadalete and Segura basins along with smaller basins of southern Iberia close to the Gibraltar Strait (Guadalhorce, Velez and Guadalfeo basins); and a third group including the Sado Basin with high statistical support (pp=1; bootstrap=100). However, phylogenetic relationships among these three groups within *S. pyrenaicus* were not solved (Fig. 3). The species *S. carolitertii* grouped together in a polytomy almost all analysed

populations from Douro Basin and other basins from the northwestern Iberian quadrant (Fig. 1), except for the Mondego Basin, which constituted an independent and highly supported clade (pp=1; bootstrap=100). One individual from one tributary of the Tagus Basin (Arrago R.) was included within the *S. carolitertii* clade (Fig. 3). One population of the Tagus Basin (Alberche R.) was related closer to *S. carolitertii* than to *S. pyrenaicus*. Furthermore, two populations, the Ompolveda R. (Upper Tagus Basin) and the Piedra R. (Ebro Basin) revealed some individuals related to *S. pyrenaicus* but some others to the species *S. castellanus* (Fig. 3).

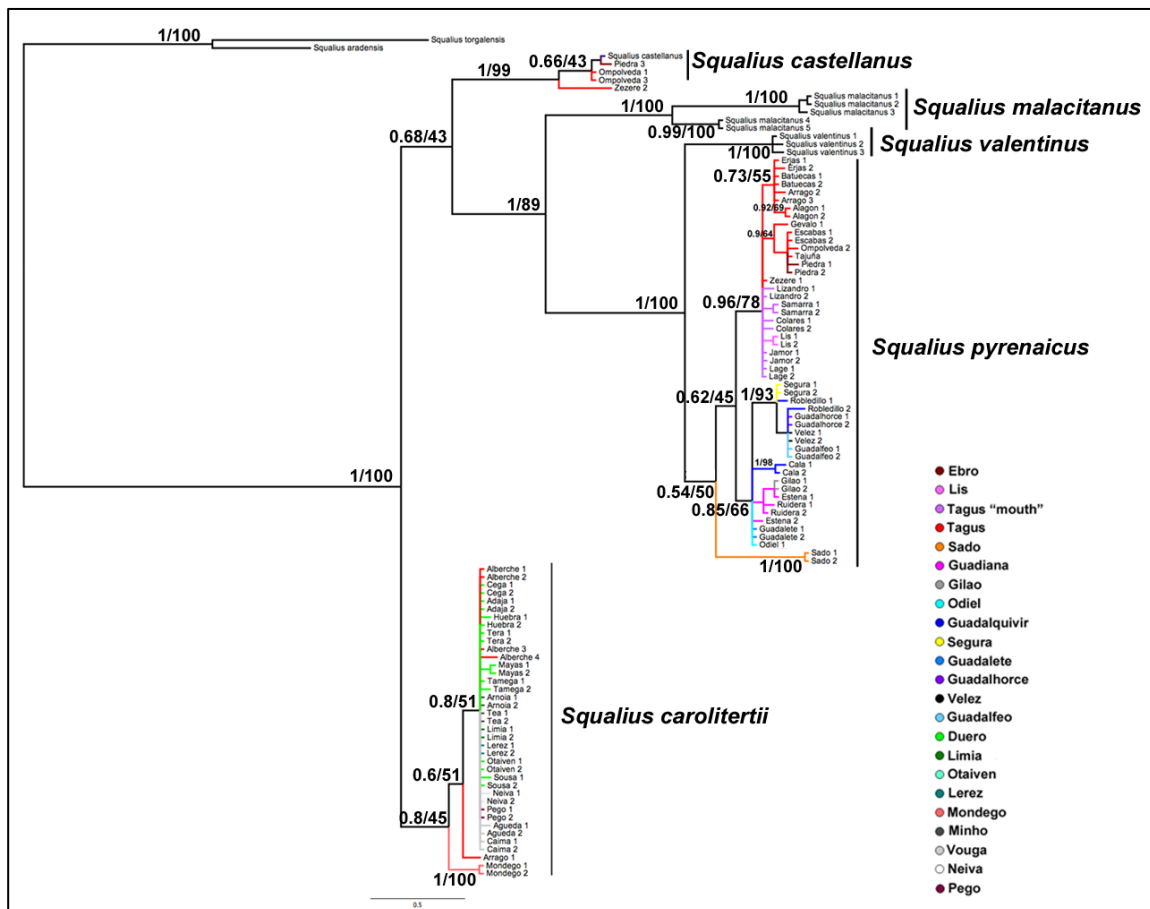


Fig. 3. Phylogenetic tree based on *MT-CYB* gene (720 bp). Bayesian Inference and Maximum Likelihood analyses. Before dash posterior probabilities of BI. After dash bootstrap values of ML. Tagus "mouth" refers to small basin drainages close to the mouth of the Tagus R.

Nuclear phylogenetic relationships of all analysed genes concatenated (4866 bp) based on BI and ML were not congruent with mitochondrial phylogeny and showed the divergence of the clade comprised of the species *S. aradensis* and *S. torgalensis* from a large clade constituted by two subclades, one moderately supported clade (pp=0.9) constituted by the species *S. castellanus* and *S. carolitertii* along with the mitochondrial lineage of *S. pyrenaicus* including

populations from the Tagus and Ebro drainages and small basins close to the Tagus mouth, and a second strongly supported subclade (pp=1; bootstrap=100) clustering the species *S. malacitanus* and *S. valentinus* together with populations of *S. pyrenaicus* from the southern half of Iberian Peninsula, although the statistical support of these three latter *S. pyrenaicus* populations was moderate (pp=0.88; bootstrap=47) (Fig. 4). Therefore, the species *S. pyrenaicus* was polyphyletic on the basis of nuclear phylogeny. In this second subclade nuclear monophyly of the species *S. malacitanus*, *S. valentinus* and *S. pyrenaicus* from Sado Basin was confirmed with high support, as well as the monophyly of *S. carolitertii* in the first subclade (Figs. 3 and 4).

Concatenated mitochondrial plus nuclear genes (5586 bp) rendered better resolved phylogenetic relationships among Iberian *Squalius* species than mitochondrial or nuclear phylogenies considered independently and reinforced the relationship of a clade constituted by the species *S. castellanus* and a divergent lineage from the Zezere R. (Tagus Basin) as sister species of *S. carolitertii* (Fig. 5; pp=1; bootstrap=100). This concatenated phylogeny also supported the inclusion of some specimens from some tributaries of the Upper Tagus and Ebro basins in the species *S. castellanus*, which showed mitochondrial genetic distances ranging from 3.6 to 3.7% with relation to *S. carolitertii* and from 5.2 to 5.6% relative to *S. pyrenaicus*. As for nuclear phylogeny, the species *S. pyrenaicus* was not recovered as a monophyletic clade due to the phylogenetic position of *S. valentinus* as sister taxa to the *S. pyrenaicus* populations from the southern half of Iberian Peninsula, although this relation of sister taxa did not show high support (pp=0.71; bootstrap=25). Nonetheless, *S. pyrenaicus* populations from Tagus, Ebro and small basins close to the Tagus mouth clustered together within a moderately supported clade (pp=0.94), relationships that were not congruent with nuclear phylogeny. Relationships within the species *S. carolitertii* and *S. pyrenaicus* were congruent with those found in only-mitochondrial topology. Mitochondrial genetic distances among and within species for both *S. carolitertii* and *S. pyrenaicus* relative to other Iberian species of this genus are represented in Tables 3, 4 and 5.

Table 3. Uncorrected-p genetic distances between Iberian *Squalius* species analysed based on MT-CYB (below branches). Uncorrected-p genetic distances within species on diagonal (in bold).

	S. castellanus	S. carolitertii	S. malacitanus	S. pyrenaicus	S. valentinus	Zezere R.
<i>S. castellanus</i>	0.002					
<i>S. carolitertii</i>	0.036	0.003				
<i>S. malacitanus</i>	0.069	0.063	0.019			
<i>S. pyrenaicus</i>	0.055	0.060	0.067	0.011		
<i>S. valentinus</i>	0.067	0.064	0.070	0.029	0.001	
Zezere R.	0.013	0.034	0.071	0.059	0.067	0.002

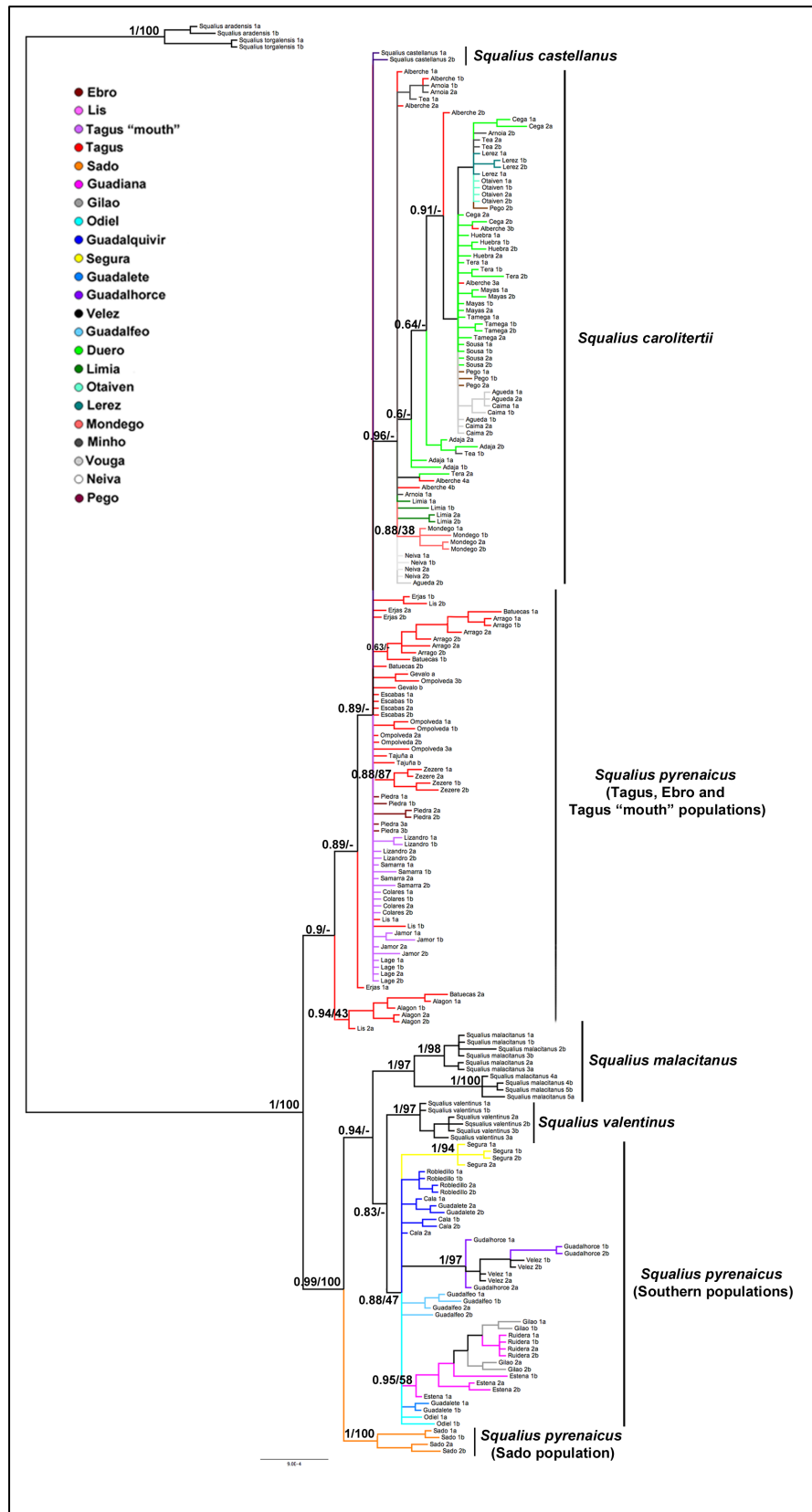


Fig. 4. Concatenated nuclear phylogenetic tree based on Bayesian Inference and Maximum Likelihood analyses and six nuclear markers (4866 bp). Before dash posterior probabilities of BI. After dash bootstrap values of ML. Tagus "mouth" refers to small basin drainages close to the mouth of the Tagus R.

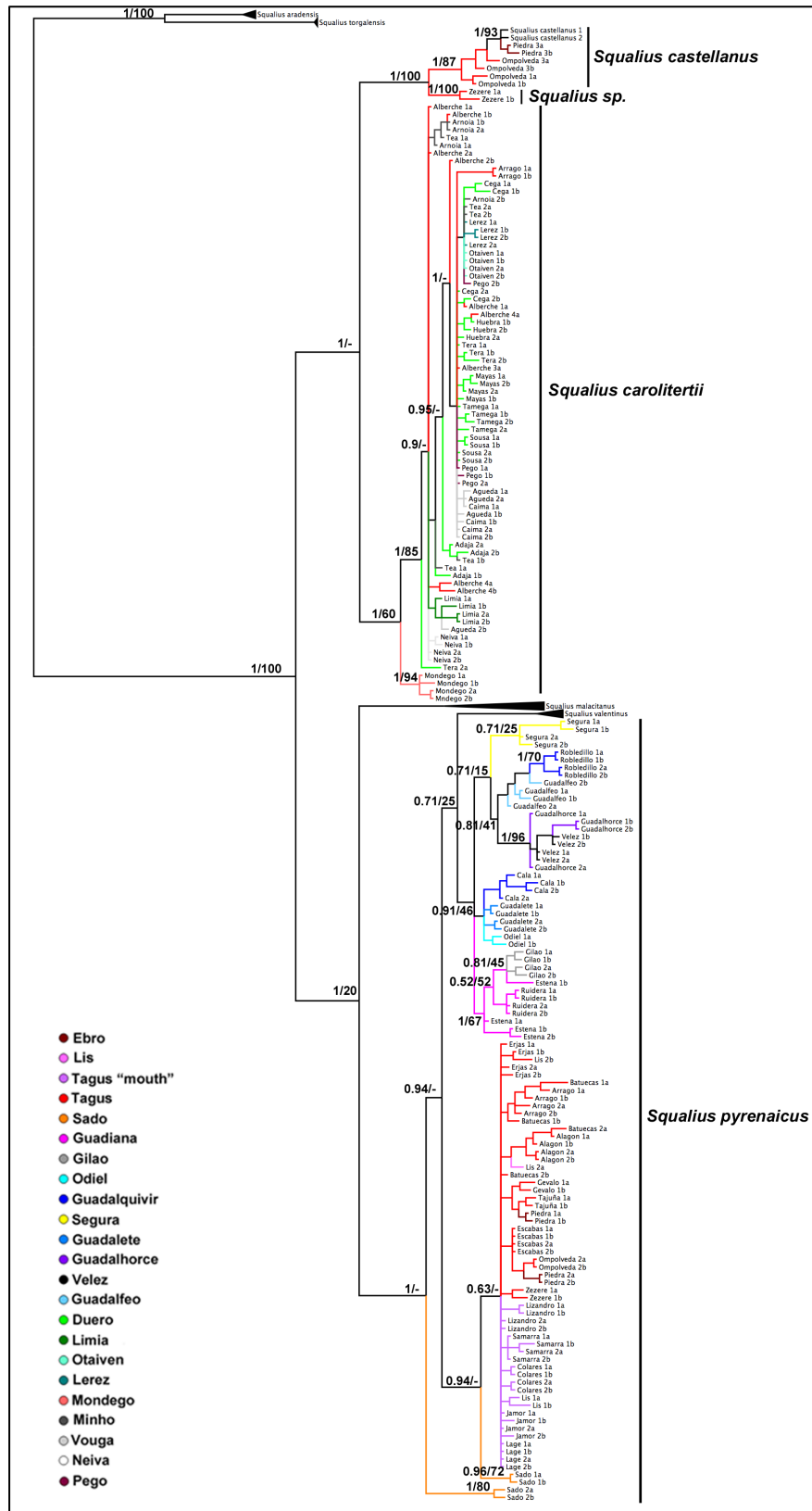


Fig. 5. Concatenated mitochondrial + nuclear phylogenetic tree based on Bayesian Inference and Maximum Likelihood analyses (5586 bp). Before dash posterior probabilities of BI. After dash bootstrap values of ML. Tagus "mouth" refers to small basin drainages close to the mouth of the Tagus R.

Table 4. Uncorrected-*p* genetic distances between *Squalius carolitertii* populations based on *MT-CYB* (below branches). Uncorrected-*p* genetic distances within populations on diagonal (in bold).

	Alcoa	Mondego	Vouga	Ave	Cavado	Neiva	Pego	Limia	Otaiven	Lerez	Minho	Douro	Alberche R.	Zezere R.
Alcoa	0.000													
Mondego	0.001	0.002												
Vouga	0.009	0.008	0.000											
Ave	0.009	0.009	0.000	0.001										
Cavado	0.009	0.009	0.001	0.001	0.001									
Neiva	0.008	0.009	0.001	0.000	0.001	0.001								
Pego	0.009	0.008	0.000	0.000	0.001	0.000	0.000							
Limia	0.009	0.009	0.001	0.001	0.001	0.001	0.000	0.001						
Otaiven	0.008	0.008	0.000	0.000	0.001	0.000	0.000	0.000	0.000					
Lerez	0.008	0.008	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000				
Minho	0.008	0.008	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000			
Douro	0.009	0.009	0.001	0.001	0.001	0.001	0.000	0.001	0.000	0.001	0.001	0.001		
Alberche R.	0.009	0.009	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.000	0.001	0.001	
Zezere R.	0.031	0.032	0.034	0.034	0.033	0.034	0.034	0.034	0.034	0.034	0.034	0.034	0.034	0.002

Table 5. Uncorrected-*p* genetic distances between Iberian *Squalius pyrenaicus* populations based on *MT-CYB* (below branches). Uncorrected-*p* genetic distances within populations on diagonal (in bold).

	Ebro	Lis	Lizandro	Samarra	Colares	Lage	Jamor	Tagus	Sado	Guadiana	Gilão	Odiel	Guadalquivir	Segura	Guadalete	Guadalhorce	Velez	Guadalfeo
Ebro	0.001																	
Lis	0.008	0.001																
Lizandro	0.007	0.003	0.001															
Samarra	0.008	0.003	0.003	0.000														
Colares	0.007	0.002	0.002	0.002	0.001													
Lage	0.006	0.003	0.001	0.002	0.001	0.000												
Jamor	0.006	0.002	0.003	0.002	0.001	0.001	0.001											
Tagus	0.008	0.005	0.005	0.005	0.004	0.003	0.004	0.005										
Sado	0.028	0.024	0.022	0.024	0.023	0.022	0.022	0.024	0.003									
Guadiana	0.015	0.011	0.010	0.011	0.010	0.010	0.009	0.011	0.023	0.004								
Gilão	0.015	0.011	0.011	0.011	0.010	0.010	0.009	0.012	0.024	0.003	0.000							
Odiel	0.013	0.009	0.009	0.009	0.008	0.008	0.007	0.014	0.026	0.007	0.007	0.004						
Guadalquivir	0.017	0.013	0.013	0.013	0.012	0.012	0.011	0.014	0.025	0.007	0.007	0.007	0.007					
Segura	0.016	0.012	0.012	0.012	0.011	0.011	0.011	0.013	0.026	0.007	0.008	0.006	0.007	0.005				
Guadalete	0.014	0.010	0.010	0.010	0.009	0.009	0.008	0.011	0.023	0.007	0.004	0.005	0.005	0.005	0.002			
Guadalhorce	0.018	0.014	0.013	0.014	0.013	0.013	0.012	0.015	0.025	0.007	0.008	0.005	0.007	0.003	0.005	0.000		
Velez	0.018	0.015	0.014	0.014	0.014	0.013	0.013	0.015	0.025	0.008	0.009	0.005	0.007	0.004	0.006	0.001	0.001	
Guadalfeo	0.018	0.018	0.018	0.018	0.017	0.017	0.017	0.016	0.026	0.008	0.009	0.006	0.008	0.004	0.006	0.001	0.002	0.002

Species tree analyses based on mitochondrial and nuclear genes and only-nucle concatenated rendered incongruent topologies, as showed a cloudogram for posterior distributions for both analyses (Fig. 6). Therefore, concatenated mitochondrial plus nuclear species tree recovered a highly congruent topology with the one obtained from phylogenetic analysis using the same data set (Fig. 7A). As well as occurred in concatenated analysis, relationships among northern and southern *Squalius pyrenaicus* lineages and *S. valentinus* were not highly supported (pp=0.5). Relationship of these latter species with *S. malacitanus* showed also low support (pp=0.6). Remaining probabilities of speciation were always higher than 0.84 (Fig. 7A). Nevertheless, species tree based on only nuclear genes (Fig. 6B) rendered a different topology grouping with high statistical support (posterior probability > 0.95) northern and southern groups of *S. pyrenaicus* in different evolutionary lineages. Therefore, the northern lineage was clustered along with the Zezere population in a clade (pp for the speciation event = 0.92) sister to *S. castellanus*, which in turn was the sister clade to the species *S. carolitertii*. On the other hand, the southern lineage of *S. pyrenaicus* appeared as the sister species of *S. valentinus* (pp=0.96). The nuclear species tree analysis also confirmed the high divergence of the third divergent group of *S. pyrenaicus* (Sado Basin), although it was recovered as sister to the lineage including the northern *S. pyrenaicus* clade, this relationship was only moderately supported (pp=0.94), and confirm the ambiguity of Sado lineage relationship found in gene trees. A

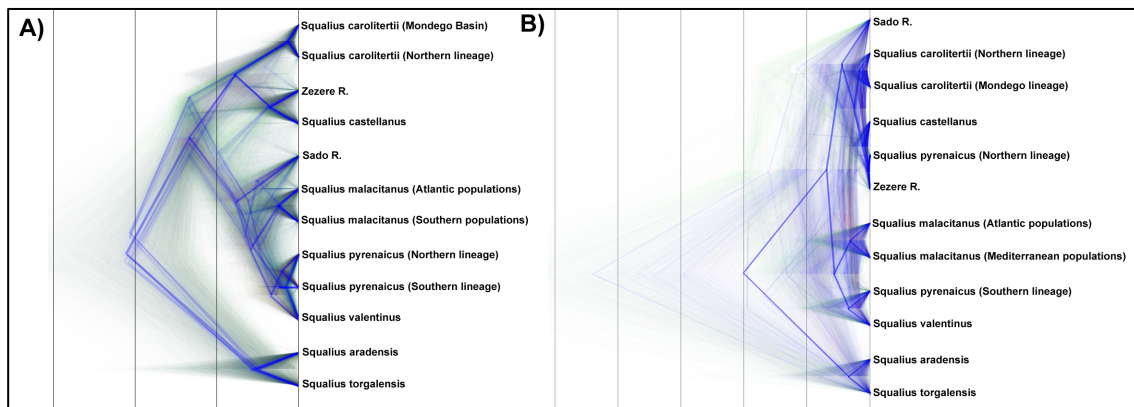


Fig. 6. Species tree consensus superimposed upon a cloudogram of the entire posterior distribution of the species trees from *BEAST analysis. A) Concatenated mitochondrial plus nuclear genes (5586 bp). B) Concatenated nuclear genes (4866 bp).

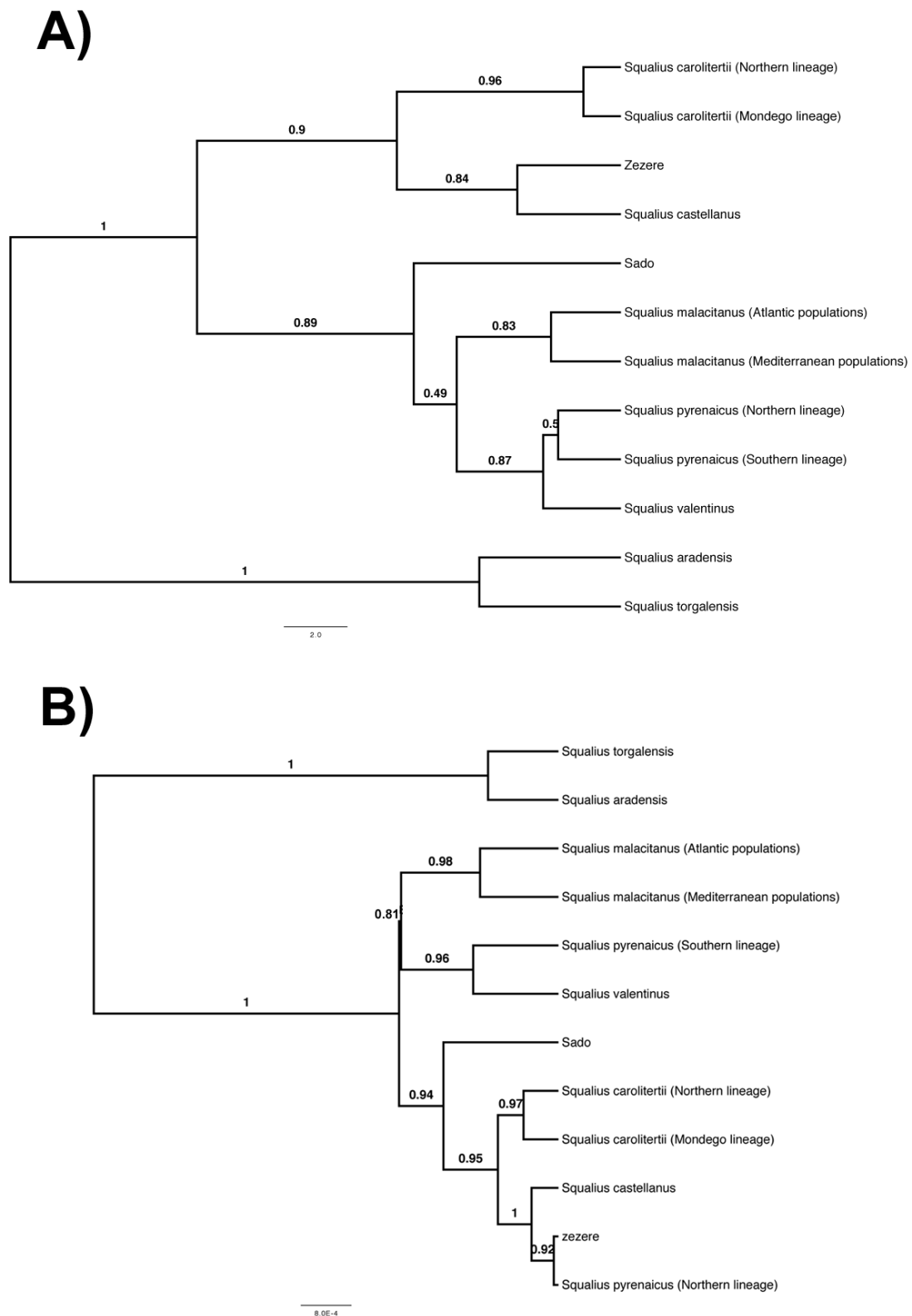


Fig. 7. Bayesian species tree delimitation for the concatenated mitochondrial and nuclear genes (A) and the concatenated nuclear genes (B). Above branches posterior probability values.

2. Phylogeography, genetic structure and genetic diversity

Haplotype networks for *MT-CYB* and *RAG1* genes including the species *S. carolitertii*, *S. castellanus*, *S. malacitanus*, *S. pyrenaicus* and *S. valentinus* underlined the complex genetic structure exhibited by *S. pyrenaicus* relative to the other Iberian species of the genus analysed, as well as the sharing of nuclear alleles between the species *S. castellanus*, *S. carolitertii* and *S. pyrenaicus* on the basis of the fragment of the nuclear marker analysed (Fig. 8A y 8B). A more specific network analysis considering only *S. carolitertii* and *S. pyrenaicus*, the two species of interest in this study, showed a clear difference in their phylogeographic structure for both analysed markers. In the mitochondrial network, the species *S. carolitertii* exhibited three different haplogroups. One star-like haplogroup containing one widespread haplotype shared by all populations of this species and several low frequent peripheral haplotypes, which erase any signal of geographic structure. A second haplogroup encompassing haplotypes from the Mondego and Alcoa basins; and a third haplogroup including samples from one Portuguese tributary of the Tagus Basin, the Zezere R., which in turn was closely related to one haplogroup that assembled together the species *S. castellanus*, with some individuals from a tributary of the Upper Tagus Basin (Ompolveda R.) and one individual from a tributary of the Ebro Basin (Piedra R.), showing twenty-one mutational steps relative to the Mondego-Alcoa clade (Figs. 2 and 9).

The species *S. pyrenaicus* showed also three independent mitochondrial lineages, but none of them displayed a clear star-like structure such was the case of *S. carolitertii*, although some haplotypes were present in high frequency in the species *S. pyrenaicus*. Instead of, a high number of low frequency haplotypes was observed (Fig. 9). The most divergent lineage of *S. pyrenaicus* (12 mutational steps) included haplotypes from the Sado Basin. A second lineage was comprised of haplotypes from Tagus Basin along with small basins close to its mouth and from two tributaries from the Ebro Basin; a partial substructure was found in this second lineage, with one subgroup including haplotypes from Lower-Middle Tagus Basin and other subgroup clustering haplotypes from the Upper Tagus and Ebro basins and a third subgroup with a highly frequent haplotype present in the small basins close to the Tagus Basin mouth (Figs. 1 and 9). The third mitochondrial lineage of the species *S. pyrenaicus* was geographically located in the southern half of Iberian Peninsula and was constituted by haplotypes from the large Guadalquivir and Guadiana basins and the smaller Gilão, Segura, Guadalete, Odiel, Guadalhorce, Velez and Guadalfeo basins (Fig. 1 and 9). A geographical substructure was also found in this third lineage, in which are significant one subgroup from Guadiana-Gilão basins and a second subgroup from southern basins close to the Gibraltar Strait (Guadalhorce, Velez and Guadalfeo) as well as some tributaries of the Upper Guadalquivir Basin. It is also noteworthy in this third lineage a small subgroup of haplotypes of one tributary of the lower part of the Guadalquivir Basin (Cala R.).

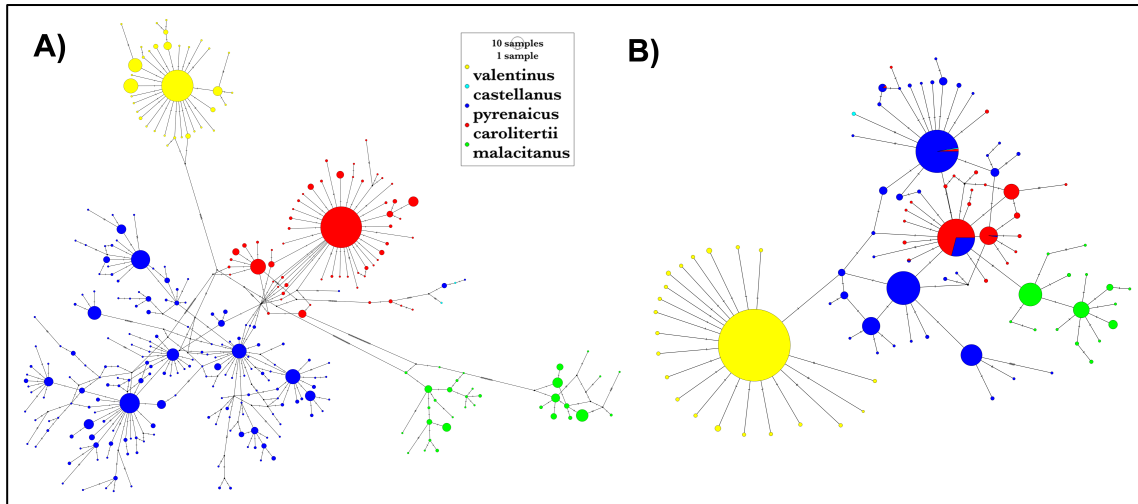


Fig. 8. Haplotype network based on *MT-CYB* gene (A; 720 bp; 2347 individuals) and *RAG1* gene (B; 832 bp; 908 individuals) belonging to several Iberian *Squalius* species (see legend).

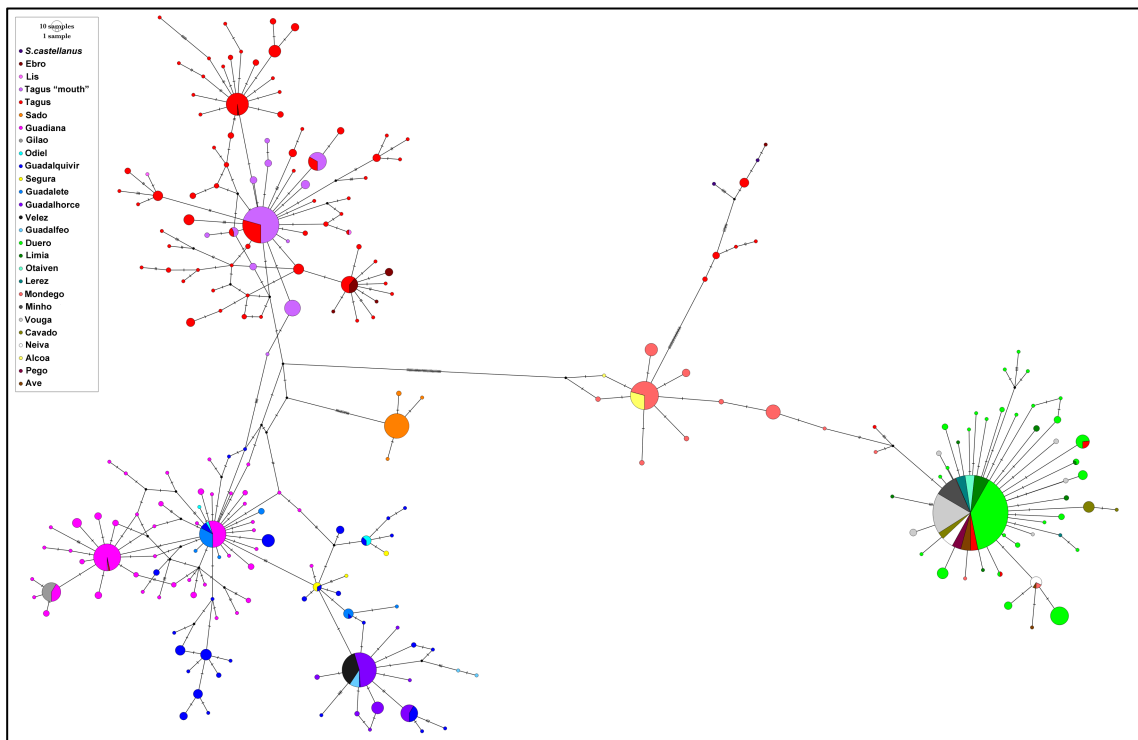


Fig. 9. Haplotype network based on *MT-CYB* gene (720 bp) and 1658 individuals belonging to the species *Squalius castellanus*, *S. carolitertii* and *S. pyrenaicus*.

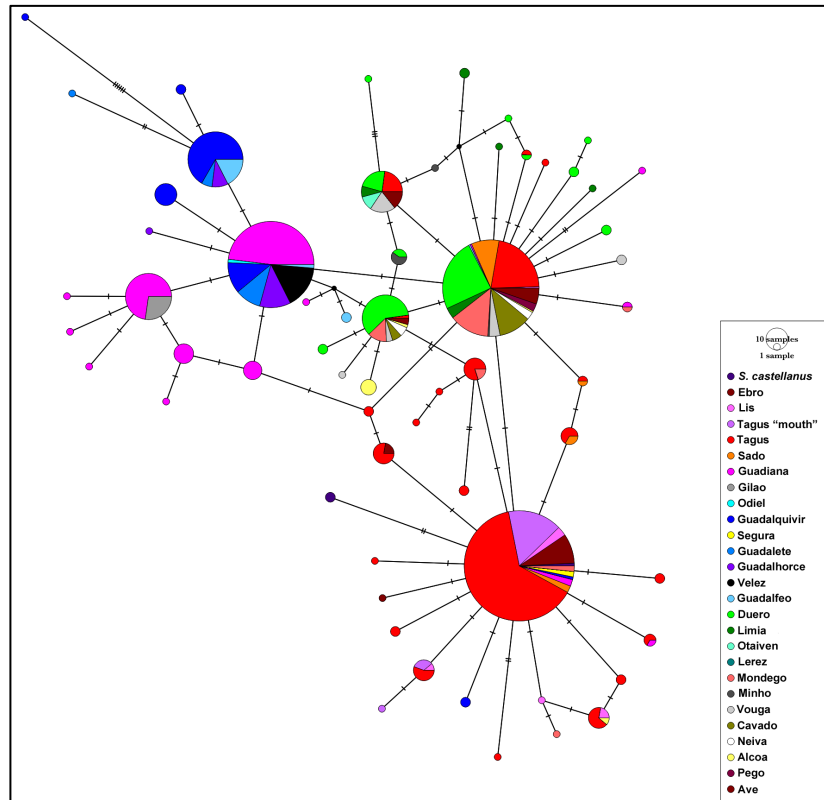


Fig. 10. Haplotype network based on *RAG1* gene (832 bp) and 456 individuals belonging to the species *Squalius castellanus*, *S. carolitertii* and *S. pyrenaicus*.

The species *S. carolitertii* showed significant genetic structure among basins, among populations within basins and within populations when AMOVA analysis considered hydrological basin structure as well as when the partitioning scheme of the genetic variance was based on mitochondrial lineages (Table 6). Nevertheless, the significant genetic structure among basins is an artefact caused by the presence of different mitochondrial lineages, as demonstrated AMOVA analyses considering populations of the mitochondrial lineage from Douro and related basins (non-significant ϕ_{CT}) and the highest percentage of genetic variance (87.10%) in the mitochondrial lineage analysis (Table 6). In the species *S. pyrenaicus*, when hierarchical genetic structure based on hydrological basins was evaluated the three source of genetic variation (among basins, among populations within basins and within basins) were significant in AMOVA analyses, showing for the *MT-CYB* gene significant genetic structure for the three partitioned components (Table 7). In this species the percentage of variation was higher among basins (56.7%).

On the other hand, ϕ_{ST} pairwise comparisons and Nm values for *S. carolitertii* and *S. pyrenaicus* are represented in Table 8 and Table 9 respectively. For *S. pyrenaicus*, ϕ_{ST} were usually higher than 0.7 in comparisons concerning different mitochondrial lineages, except for the

Tagus population, although values in pairwise comparisons with populations belonging to its mitochondrial lineage were lower than those comparisons with individuals from different lineages. Besides ϕ_{ST} comparisons of involving the Sado Basin were always higher than 0.9.

Tabla 6. AMOVA performance for the species *S. carolitertii* based on the *MT-CYB* gene.

Structure tested	d.f.	SSD	Variance components	% genetic variance	ϕ statistic	p-value
Hidrological basins						
Among basins	12	520.447	0.79487	64.12	$\phi_{CT} = 0.64116$	< 0.0001
Among populations within basins	41	99.877	0.20968	16.91	$\phi_{SC} = 0.47133$	< 0.0001
Within populations	664	156.168	0.23519	18.97	$\phi_{ST} = 0.81029$	< 0.0001
Hidrological basins (without Mondego and Alcoa basins)						
Among basins	8	10.899	-0.03899	-12.98	$\phi_{CT} = -0.1297$	0.85337
Among populations within basins	38	66.410	0.14623	48.67	$\phi_{SC} = 0.43076$	< 0.0001
Within populations	556	107.438	0.19323	64.31	$\phi_{ST} = 0.35690$	< 0.0001
Mitochondrial genetic lineages						
Among lineages	1	506.128	2.60384	87.10	$\phi_{CT} = 0.87098$	< 0.0001
Among populations within lineages	52	114.196	0.15051	5.03	$\phi_{SC} = 0.39023$	< 0.0001
Within populations	664	56.168	0.23519	7.87	$\phi_{ST} = 0.92133$	< 0.0001

Tabla 7. AMOVA performance for the species *S. pyrenaicus* based on the *MT-CYB* gene.

Structure tested	d.f.	SSD	Variance components	% genetic variance	ϕ statistic	p-value
Hidrological basins						
Among basins	17	2366.709	2.34592	56.69	$\phi_{CT} = 0.56692$	< 0.0001
Among populations within basins	94	614.148	0.83961	20.29	$\phi_{SC} = 0.46851$	< 0.0001
Within populations	906	862.956	0.95249	23.02	$\phi_{ST} = 0.76982$	< 0.0001
Mitochondrial genetic lineages						
Among lineages	2	1864.207	3.38324	62.40	$\phi_{CT} = 0.62399$	< 0.0001
Among populations within lineages	15	467.148	0.73828	13.62	$\phi_{SC} = 0.03613$	< 0.0001
Within populations	901	1171.695	1.30044	23.98	$\phi_{ST} = 0.76015$	< 0.0001

Table 8 . ϕ_{ST} pairwise comparisons (below diagonal) and gene flow (above diagonal) for the MT-CYB genet between populations of *S. pyrenaicus*. Gene flow is estimated as number of migrants contributing alleles each generation (Nm). Nm>1 in bold.

	Alcoa	Mondego	Vouga	Ave	Cavado	Neiva	Pego	Limia	Otaiven	Lerez	Minho	Douro	Alberche
Alcoa	-												
Mondego	0.065	-											
Vouga	0.965	0.836	-										
Ave	0.969	0.773	0.030	-									
Cavado	0.943	0.777	0.381	0.293	-								
Neiva	0.986	0.784	0.189	0.041	0.336	-							
Pego	0.991	0.774	0.016	0.024	0.340	0.180	-						
Limia	0.934	0.785	0.027	0.012	0.114	0.010	0.006	-					
Otaiven	0.992	0.777	0.012	0.030	0.352	0.189	0.000	0.006	-				
Lerez	0.984	0.776	0.002	0.023	0.325	0.170	0.005	0.003	0.000	-			
Minho	0.995	0.806	0.003	0.853	0.458	0.267	0.000	0.017	0.000	0.039	-		
Douro	0.869	0.815	0.033	0.020	0.181	0.020	0.002	0.030	0.006	0.007	0.020	-	
Alberche	0.946	0.775	0.159	0.222	0.283	0.190	0.128	0.083	0.138	0.121	0.020	0.050	

Table 9. ϕ_{ST} pairwise comparisons (below diagonal) and gene flow (above diagonal) for the *MT-CYB* gene between populations of *S. pyrenaicus*. Gene flow is estimated as number of migrants contributing alleles each generation (Nm). $Nm > 1$ in bold.

	Ebro	Lis	Lizandro	Samarra	Colares	Lage	Jamor	Tagus	Sado	Guadiana	Gilão	Odiel	Guadalquivir	Segura	Guadalete	Guadalhorce	Velez	Guadalfeo
Ebro	-	0.160	0.266	0.087	0.301	0.155	0.216	1.463	0.050	0.168	0.028	0.092	0.301	0.108	0.114	0.017	0.028	0.055
Lis	0.758	-	0.517	0.104	0.578	0.271	0.400	2.390	0.050	0.192	0.024	0.084	0.336	0.094	0.120	0.015	0.027	0.049
Lizandro	0.652	0.492	-	0.322	4.359	14.22	7.650	8.667	0.052	0.218	0.051	0.098	0.355	0.109	0.139	0.029	0.040	0.062
Samarra	0.851	0.828	0.608	-	0.338	0.068	0.166	1.723	0.042	0.182	0.003	0.051	0.319	0.054	0.096	0.008	0.018	0.027
Colares	0.624	0.464	0.103	0.596	-	6.699	14.230	9.024	0.055	0.238	0.056	0.120	0.394	0.136	0.162	0.029	0.042	0.073
Lage	0.763	0.648	0.034	0.879	0.070	-	37.50	16.58	0.048	0.217	0.007	0.067	0.377	0.073	0.120	0.022	0.010	0.035
Jamor	0.698	0.555	0.061	0.750	0.034	0.013	-	16.40	0.051	0.247	0.025	0.096	0.418	0.108	0.150	0.016	0.030	0.027
Tagus	0.255	0.173	0.054	0.225	0.052	0.029	0.029	-	0.135	0.344	0.292	0.319	0.394	0.368	0.389	0.198	0.205	0.236
Sado	0.910	0.911	0.906	0.994	0.901	0.913	0.906	0.788	-	0.079	0.043	0.060	0.079	0.062	0.063	0.030	0.036	0.054
Guadiana	0.748	0.722	0.696	0.733	0.677	0.697	0.904	0.592	0.863	-	0.079	1.125	1.712	0.522	2.516	0.248	0.261	0.313
Gilão	0.946	0.955	0.908	0.921	0.899	0.986	0.953	0.631	0.921	0.307	-	0.110	0.796	0.092	0.279	0.011	0.029	0.045
Odiel	0.844	0.856	0.836	0.907	0.806	0.881	0.839	0.610	0.892	0.435	0.350	-	2.970	2.016	0.929	0.110	0.177	0.488
Guadalquivir	0.624	0.598	0.584	0.610	0.559	0.570	0.823	0.559	0.787	0.226	0.887	0.144	-	2.704	4.522	0.817	0.854	1.180
Segura	0.822	0.842	0.820	0.902	0.786	0.872	0.968	0.576	0.890	0.489	0.386	0.199	0.890	-	0.796	0.247	0.401	1.339
Guadalete	0.814	0.806	0.782	0.839	0.755	0.806	0.769	0.562	0.887	0.166	0.641	0.350	0.887	0.386	-	0.144	0.185	0.321
Guadalhorce	0.967	0.970	0.946	0.985	0.945	0.981	0.943	0.716	0.943	0.668	0.776	0.820	0.943	0.669	0.776	-	2.352	1.527
Velez	0.947	0.948	0.927	0.965	0.922	0.958	0.669	0.709	0.932	0.657	0.729	0.739	0.932	0.555	0.729	0.175	-	1.851
Guadalfeo	0.900	0.911	0.889	0.949	0.872	0.934	0.544	0.679	0.902	0.615	0.607	0.506	0.902	0.272	0.607	0.247	0.213	-

Tabla 10. Genetic diversity statistics estimated in *S. carolitertii* and *S. pyrenaicus* for the *MT-CYB* gene.

Population	Number of individuals analyzed	Number of haplotypes	Haplotype diversity (SD)	Nucleotide diversity (SD)	K	S
<i>Squalius carolitertii</i>						
Alcoa	20	2	0.100 (0.088)	0.00014 (0.00012)	0.100	1
Mondego	95	13	0.717 (0.04)	0.00243 (0.00035)	1.750	14
Vouga	92	6	0.224 (0.058)	0.00033 (0.00009)	0.235	5
Ave	21	3	0.186 (0.110)	0.00038 (0.0004)	0.276	2
Cavado	25	3	0.547 (0.054)	0.00082 (0.00012)	0.593	2
Neiva	31	2	0.396 (0.078)	0.00055 (0.00011)	0.396	1
Pego	18	1	0.000 (0.000)	0.000 (0.000)	0.000	0
Limia	41	7	0.390 (0.095)	0.0008 (0.00026)	0.576	9
Otaiven	20	1	0.000 (0.000)	0.000 (0.000)	0.000	0
Lerez	20	2	0.100 (0.088)	0.00014 (0.00012)	0.100	1
Minho	43	1	0.000 (0.000)	0.000 (0.000)	0.000	0
Douro	271	29	0.551 (0.0013)	0.00132 (0.00012)	0.950	37
Alberche R.	21	3	0.452 (0.105)	0.00079 (0.00026)	0.571	3
<i>Squalius pyrenaicus</i>						
Ebro	17	5	0.662 (0.094)	0.00143 (0.00044)	1.029	6
Lis	20	3	0.195 (0.115)	0.00096 (0.00058)	0.689	6
Lizandro	38	6	0.586 (0.087)	0.00141 (0.00027)	1.013	5
Samarra	22	2	0.091 (0.081)	0.00013 (0.00011)	0.091	1
Colares	26	4	0.612 (0.081)	0.001 (0.00018)	0.717	3
Lage	20	2	0.189 (0.108)	0.00026 (0.00015)	0.189	1
Jamor	20	2	0.100 (0.088)	0.00083 (0.00073)	0.600	6
Tagus	239	72	0.944 (0.008)	0.00536 (0.00032)	3.858	99
Sado	56	6	0.235 (0.075)	0.00259 (0.00125)	1.864	20
Guadiana	171	43	0.847 (0.021)	0.00354 (0.00031)	2.481	53
Gilão	17	1	0.000 (0.000)	0.000 (0.000)	0.000	0
Odiel	8	3	0.607 (0.164)	0.00407 (0.00098)	2.929	6
Guadalquivir	108	35	0.939 (0.010)	0.00705 (0.00028)	5.077	37
Segura	6	3	0.600 (0.215)	0.00509 (0.00265)	3.667	11
Guadalete	32	6	0.607 (0.081)	0.00265 (0.00045)	1.911	8
Guadalhorce	50	4	0.154 (0.068)	0.00022 (0.00010)	0.158	3
Velez	49	4	0.438 (0.0044)	0.00066 (0.00012)	0.478	2
Guadalfeo	11	3	0.345 (0.172)	0.00207 (0.00099)	1.491	5

For the species *S. carolitertii*, the highest mitochondrial genetic diversity values (haplotype and nucleotide diversity) were found in the Mondego Basin, followed by the Douro and Cavado basins (Table 10). On the contrary, the lowest diversity values were identified in the small-medium sized basins northern to the Douro Basin, which only showed one mitochondrial

haplotype in spite of the high number of individuals analysed (Tabla 10). For the species *S. pyrenaicus*, genetic diversity values were in general higher than values in *S. carolitertii*. The highest genetic diversity values were found in the largest basins, the Tagus, the Guadalquivir and the Guadiana, whereas the lowest values were estimated in small basins close to the Tagus mouth and the Gilão Basin (Table10).

3. Divergence time estimation and historical biogeography

Regarding divergence time estimations based on the concatenated mitochondrial plus nuclear species tree (Fig. 11), the separation of the clade constituted by the lineages of *S. carolitertii*, *S. castellanus* and the Mondego and Zezere basins from the lineages of *S. malacitanus*, *S. pyrenaicus*, *S. valentinus* and the Sado population took place in the Middle Miocene approximately 13.4 Ma (95% HPD: 11.4-15.5). The divergence of the species *S. carolitertii* and Mondego Basin from the clade comprised of the species *S. castellanus* plus Zezere Basin took place in the Upper Miocene (7.1 Ma; 95% HPD: 2.3-12.5 Ma), as well as the divergence of Sado population from the clade grouping *S. malacitanus*, *S. pyrenaicus* and *S. valentinus* (6.7 Ma; 95% HPD: 3.2-10.1 Ma). Pliocene divergences isolated these three later lineages, although this divergence has to be taking with caution due to the low support of these clades as consequence of mitonuclear discordance of *S. pyrenaicus* relationships, thus this divergence pattern is not reflecting the true history of the speciation process but the maternal history. The population from the Zezere Basin also diverged from the species *S. castellanus* in Pliocene period (3.3 Ma; 95% HPD: 0.5-7.3 Ma), although posterior probability supporting this divergence was low (pp=0.84).

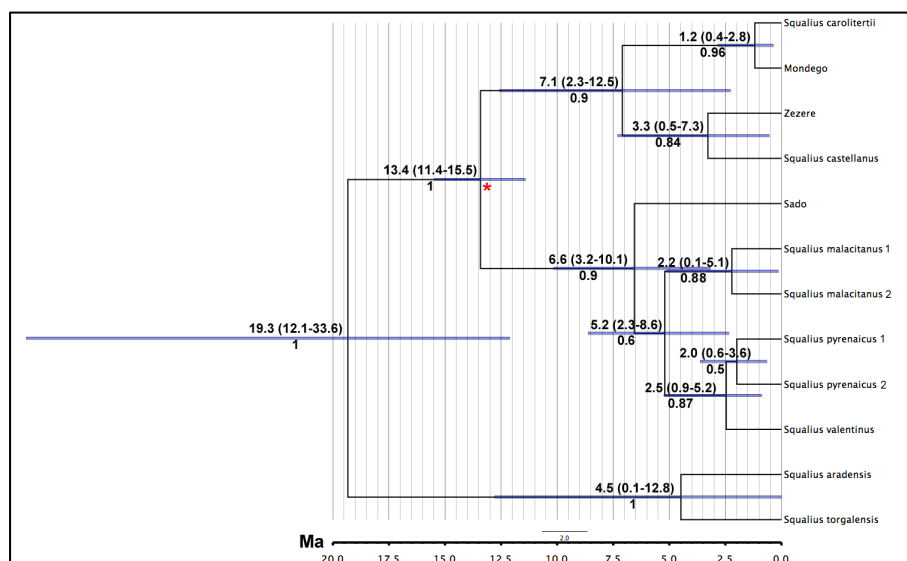


Fig. 11. Species tree and chronogram based on one mitochondrial and six nuclear genes (5586 bp and 105 individuals) obtained from *BEAST. *Squalius pyrenaicus* 1 means populations from Tagus, Ebro and small basins close to the Tagus mouth. *Squalius pyrenaicus* 2 means populations from the southern Iberian half. *

red I) means fossil calibration (*Leuciscus antunesi*: 13.5-14.5 Ma)

In spite of the mitonuclear discordance found, a more detailed divergence analysis of Iberian *Squalius* populations was taken from the mitochondrial gene tree obtained from the same *BEAST analysis (Fig. 12) in order to reconstruct the biogeographical patterns of the species *S. carolitertii* and in the northern and southern lineages of *S. pyrenaicus*, as mitochondrial marker resolved better than nuclear genes intraspecific relationships, thus allowing reconstructing biogeographical patterns among populations. In this chronogram, tMRCA of the main nodes were slightly older than the estimates obtained for the speciation events in the mitochondrial plus nuclear concatenated species tree, probably as consequence of a higher evolutionary rate of the mitochondrial genome and a reliable adjustment of the different evolutionary rates of all used markers in the species tree relative to the concatenated gene tree, although 95% HPD intervals were congruent in both chronograms. This analysis also showed that diversification within populations took place in the Pliocene and Pleistocene periods and that the most recent diversification occurred in *S. carolitertii* from Douro and other smaller basins northern to Douro (tMRCA < 2.5 Ma) (Fig. 12).

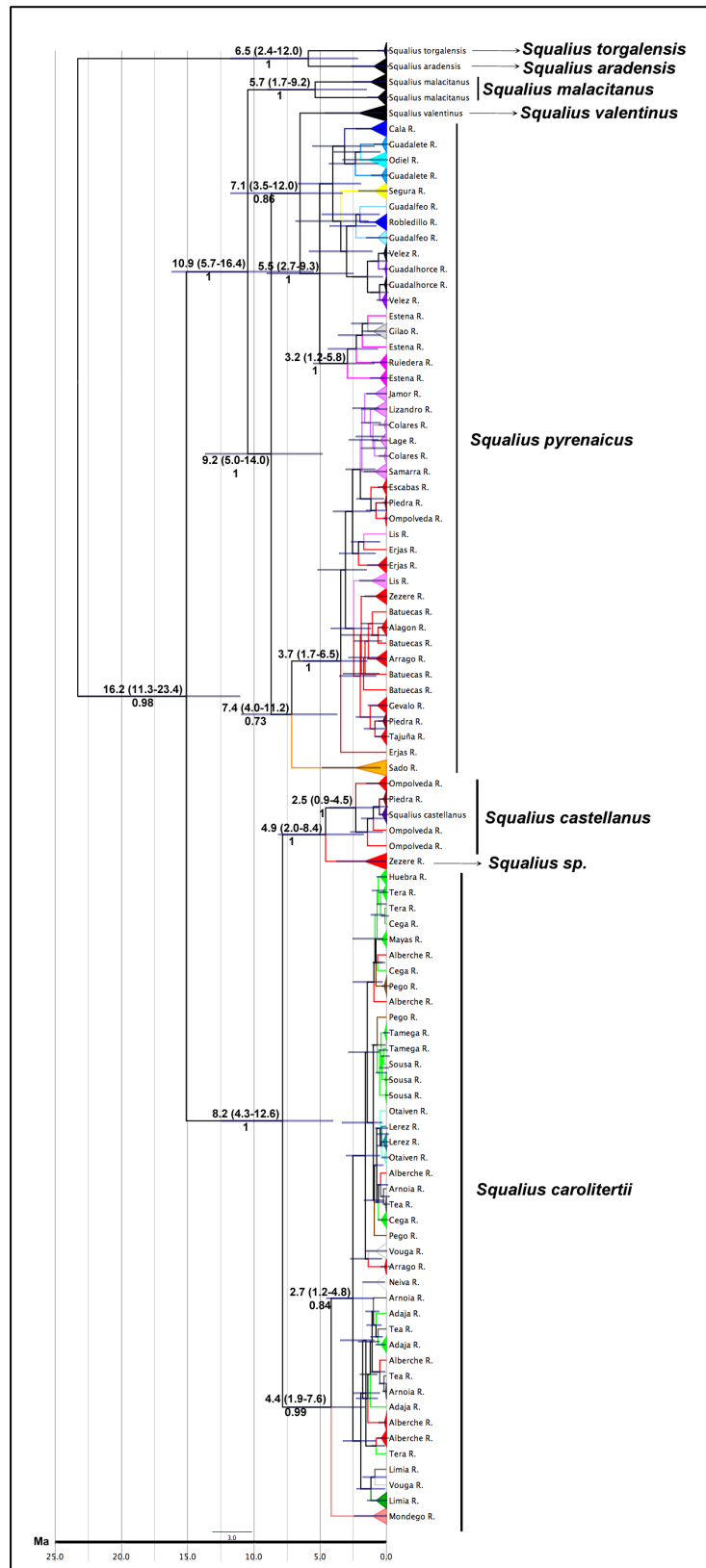


Fig. 12. Chronogram of mitochondrial gene tree (720 bp and 105 individuals) obtained from BEAST analysis after running the multispecies coalescent model using *BEAST for the concatenated mitochondrial + nuclear data matrix.

Therefore, BayArea analysis was based on the mitochondrial gene tree obtained in *BEAST analysis for all genes displaying all analysed individuals for this data set (105) and allowing the inference of historical biogeographic patterns within these three lineages of Iberian *Squalius* (Fig. 10). In the species *Squalius carolitertii*, a vicariance between the Mondego Basin and the remaining basins occurred, isolating the Mondego Basin in the Upper Miocene (Fig 13; marginal probability = 0.49). Nevertheless, reticulated phylogenetic patterns within *S. carolitertii*, not showing monophyletic populations, make difficult the inference of more specific biogeographic patterns within this species. Only some clear biogeographic patterns may be inferred, such is the case of a dispersal event from Douro to Lerez and Otaiven basins and posterior vicariance between these two latter basins and Douro (Fig. 13; marginal posterior probability = 0.65), as well as from Douro to Neiva Basin (Fig. 13; marginal probability = 0.53).

Within the species *Squalius pyrenaicus*, its northern lineage showed several vicariant and dispersal events involving the Tagus Basin and an ancestral range sharing the Tagus Basin and the small basins close to its mouth in the Upper Pliocene period. Dispersal (Fig. 14; marginal probability = 0.48) and later vicariance between the Tagus and the Lis basins (Fig. 14; marginal probability = 0.62), on one side, and between the Tagus Basin and other small basins close to its mouth (marginal probability = 0.40) on the other side were inferred to take place in the Pleistocene period.

Within the *S. pyrenaicus* populations from the southern half of Iberian Peninsula, a first vicariance between the Guadiana and Gilão basins from the rest of the basins was inferred to take place in the Upper Miocene – Lower Pliocene (Fig. 15. Marginal probability = 0.10) as well as between the Segura Basin and an area shared by the tributaries of the Upper Guadalquivir and Guadalfeo basins (marginal probability = 0.12). Furthermore, a dispersal event from a shared area between tributaries of the Upper part of the Guadalquivir Basin and Guadalfeo to the Guadalhorce and Velez basins during the Upper Pliocene - Pleistocene and a later vicariance between these two latter basins close to the Gibraltar Strait and the Upper Guadalquivir plus Guadalfeo area was estimated (Fig. 15; marginal probability = 0.2). A similar Upper Pliocene biogeographic pattern was identified in the lower part of the Guadalquivir Basin, with a vicariance between a tributary of the Lower Guadalquivir Basin (Cala R.) and the Guadalete Basin (Fig. 15; marginal probability = 0.11) followed by a dispersal event from Guadalete to Odiel Basin and posterior vicariance between these latter two basins (Fig. 15; marginal probability = 0.26).

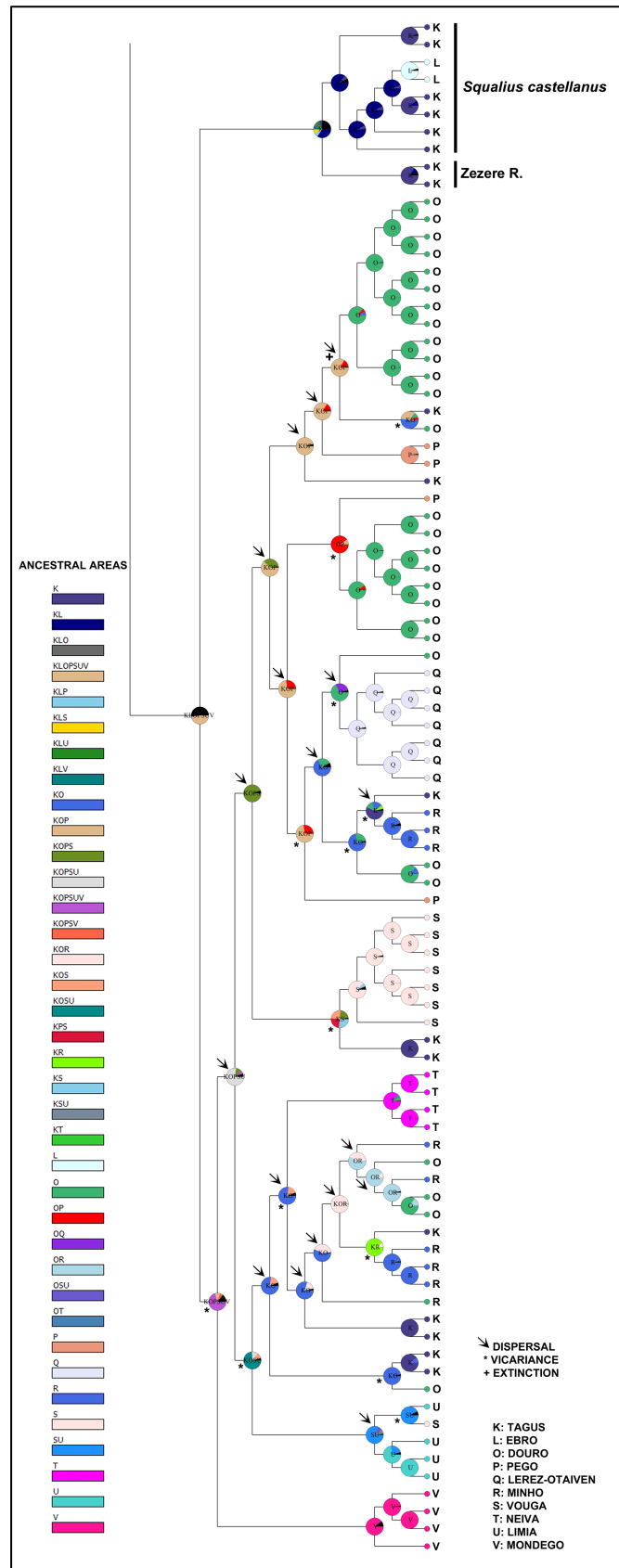


FIG. 13. Ancestral area reconstruction of *S. carolitertii* based on BayArea method implemented in RASP software.

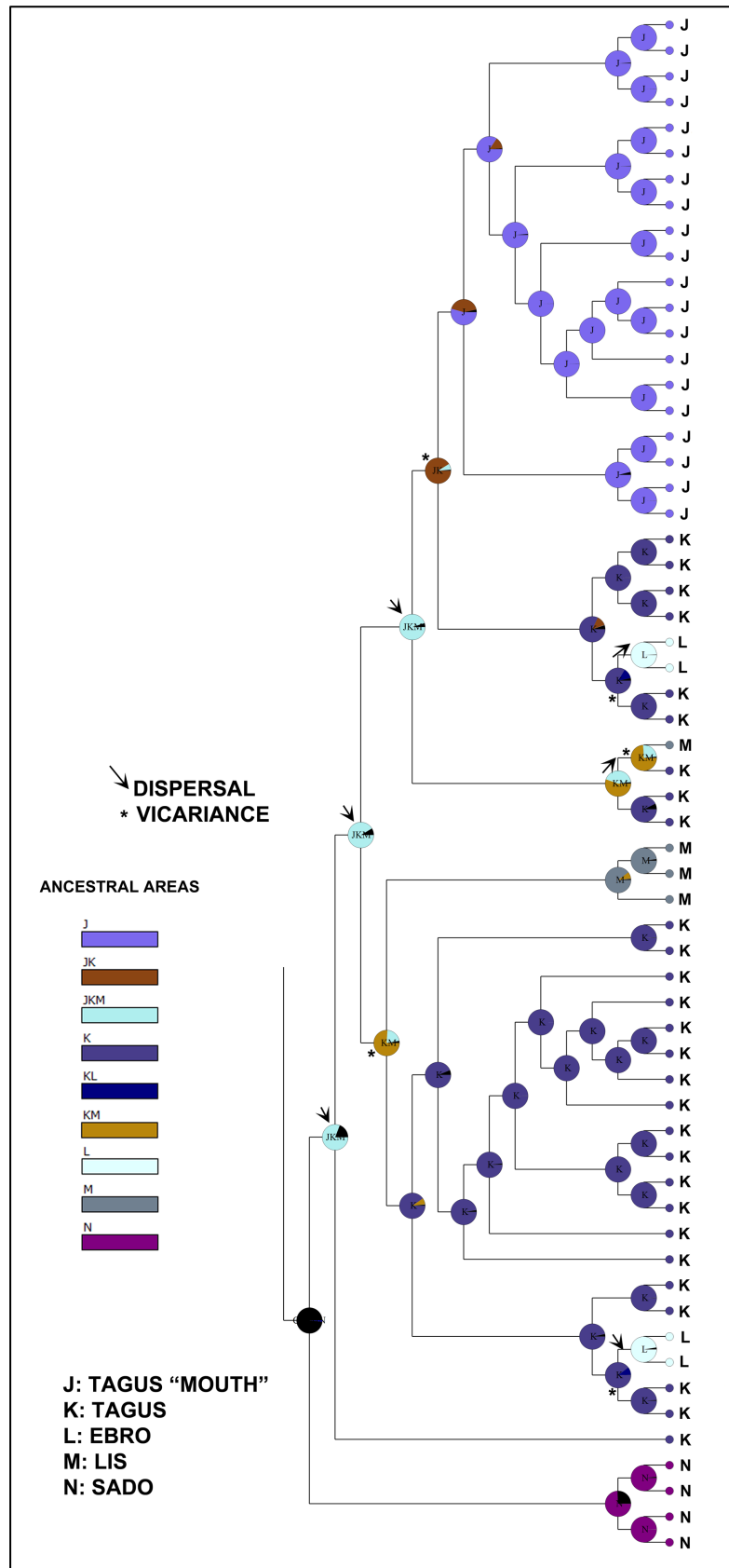


FIG. 14. Ancestral area reconstruction of the northern lineage of *S. pyrenaicus* based on BayArea method implemented in RASP software.

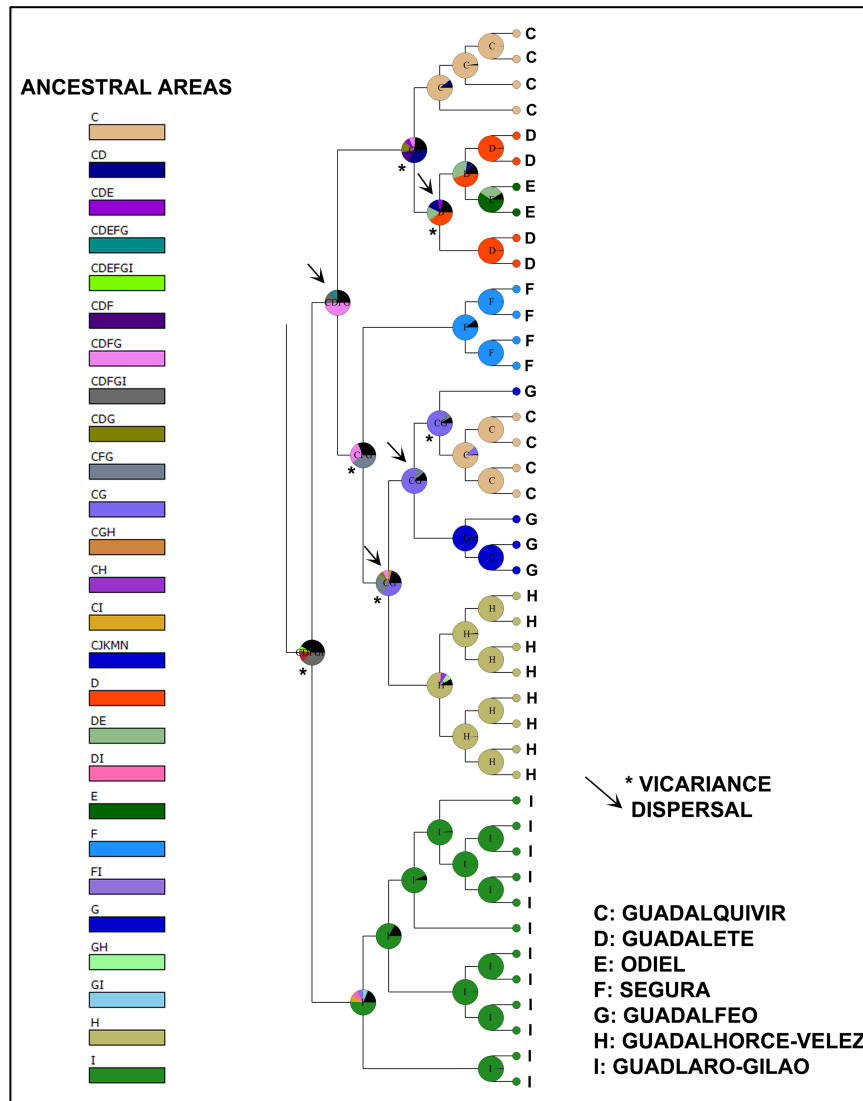


FIG. 15. Ancestral area reconstruction of *S. pyrenaicus* based on BayArea method implemented in RASP software.

DISCUSSION

In spite of the species *S. carolitertii* and *S. pyrenaicus* are the ones with the wider distribution range among the Iberian species of this genus their phylogenetic and phylogeographical relationships at intraspecific level encompassing their whole distribution range are not well established yet, as well as occur with their biogeographical history. The genetic diversity patterns of both species are not well known either, being these two issues the main objectives that we addressed in this study. The evolutionary history of *S. carolitertii* and *S. pyrenaicus* in the Iberian Peninsula is old, as have been broadly demonstrated in previous studies (Doadrio *et al.*, 2003, 2004; Sanjur *et al.*, 2003; Sousa-Santos *et al.*, 2007; Almada and Sousa-Santos, 2010; Perea *et al.*, 2016). Nevertheless, the evolutionary mechanisms responsible for the phylogeographical and

genetic diversity patterns we currently found in each species differ and explain such divergent evolutionary histories in both species.

1. Phylogenetic and phylogeographic relationships within *S. carolitertii* and *S. pyrenaicus*.

Despite nuclear concatenated analyses were not able to recover *S. pyrenaicus* and *S. castellanus* as monophyletic and both species were grouped in a polytomy together with the highly supported clade of *S. carolitertii* (pp=0.96), our species tree nuclear phylogeny validated the identity for most of the Iberian *Squalius* species analysed (Coelho *et al.*, 1998; Doadrio, 1988; Doadrio and Carmona, 2006; Doadrio *et al.*, 2007). The only discordance with taxonomy of Iberian *Squalius* species found was the polyphyletic status of the species *S. pyrenaicus* as consequence of the nuclear phylogenetic position of the three divergent groups of this species. This distinction of *S. pyrenaicus* into three divergent genetic groups was already proposed on the basis of multilocus phylogenies including some Portuguese populations (Almada and Sousa-Santos, 2010; Waap *et al.*, 2010) and is corroborated here when populations from the complete distribution range of the species are analysed. One of these *S. pyrenaicus* lineages comprised populations from Tagus and small basins close to its mouth along with partial areas of the Ebro Basin (northern lineage). Populations from the southern half of the Iberian Peninsula constitute the second *S. pyrenaicus* lineage (southern lineage). The third divergent lineage was the Sado population (Sado lineage). However, mitochondrial and concatenated (mitochondrial plus nuclear) phylogenies differs from nuclear species tree analysis, grouping northern and southern lineages of *S. pyrenaicus* within the same clade, although support for these relationships was always low. Nonetheless the mitochondrial differentiation of the Sado lineage of *S. pyrenaicus* was verified in this study (Doadrio y Carmona, 2003; Sanjur *et al.*, 2003; Sousa-Santos *et al.*, 2007; Almada y Sousa-Santos, 2010; Waap *et al.*, 2010).

According to gene tree topologies considered independently, nuclear relationships were not congruent with mitochondrial-only topology due to the close relationship of the lineage of *S. pyrenaicus* from Tagus and Ebro basins with the lineage of the species *S. carolitertii* and *S. castellanus*, although this nuclear relationship was moderately supported (pp=0.9) and not recovered when the concatenated mitochondrial plus nuclear analyses was carried out. Species tree topology for concatenated genes including the mitochondrial marker did not support the nuclear relationship found either. However, the mitochondrial and nuclear monophyly of the species *S. carolitertii* was always recovered with high support, and its sister relationship with the lineage of *S. castellanus* was corroborated when all genes were considered in the analysis, which is congruent with previous studies (Doadrio *et al.*, 2007; Perea *et al.*, 2010; Perea *et al.*, 2016). Furthermore, two divergent phylogenetic lineages were also identified in the species *S.*

carolitertii, a northern lineage encompassing the Douro and Minho basins and other small basins from the northwestern Iberian quadrant and a Mondego-Alcoa lineage, congruent with other previous studies (Almada and Sousa-Santos, 2010; Waap *et al.*, 2010). The mitochondrial and nuclear distinction of these two lineages was evident and highly supported in all performed analyses. These results suggest that northern lineage of *S. pyrenaicus* and Zezere populations probably belong to the same taxa and that southern and Sado lineages possibly constitute different species, as well as the Mondego-Alcoa lineage of the species *S. carolitertii*. However, while genetic distances between Sado and the remaining *S. pyrenaicus* populations ranging from 2.2-2.8%, these values were always lesser than 1% in the comparisons between the Mondego-Alcoa lineages and *S. carolitertii* populations. Therefore, further specific studies considering the morphology of all these populations should be performed to clarify the taxonomic status of Iberian *Squalius* species, especially imperative for *S. pyrenaicus*.

The discordance between mitochondrial and nuclear data suggest a past hybridization event that resulted in a mitochondrial introgression from the southern lineage of *S. pyrenaicus* to the northern lineage, as is also revealed by the presence of shared mitochondrial haplotypes in the haplotype network, although in very low frequency, between these two lineages. Outcomes from divergence time estimates based on concatenated mitochondrial plus nuclear data suggest that mitochondrial introgression was old, as divergence between northern and southern mitochondrial lineages took place in the Upper Pliocene – Lower Pleistocene. In the Zezere population, a tributary from the lower Tagus Basin in Portugal, two different mitochondrial haplotypes were present, the ones similar to those from other regions of the Tagus Basin, but also a different mitochondrial haplotype related to *S. castellanus* and *S. carolitertii* lineage. The special geographical configuration of Zezere River, flowing by a glacial valley, should diminish the magnitude of the introgression event, avoiding the expansion of the “invader” mitochondrial genome as the same way that in other parts of the Tagus Basin. Therefore, in the Zezere Basin are present the original mitochondrial genome from the northern lineage of *S. pyrenaicus*. Sharing of mitochondrial DNA between *S. pyrenaicus* populations from Guadiana and Tagus basins had been interpreted as an introgressive hybridization as consequence of secondary river contacts in Plio-Pleistocene periods (Sousa-Santos *et al.*, 2007; Waap *et al.*, 2010) and here we provide a broad phylogenetic context to explain this event.

Numerous studies demonstrated that hybridization events have been very common in the Iberian Peninsula, probably favoured by the complex paleogeography of this region strongly affecting the reorganization of hydrological basins during Plio-Pleistocene periods (de Vicente *et al.*, 2011; Pais *et al.*, 2012). Indeed, one of the most complex and interesting cases of hybridization found in the Iberian Peninsula is that involving the species *S. alburnoides* (e.g.

Carmona and Doadrio, 1999; Cunha *et al.*, 2004; Sousa-Santos *et al.*, 2007). In our study a possible case of hybridization between *S. carolitertii* and *S. pyrenaicus* could be identified in one analysed individual from a tributary of the Tagus Basin (Arrago R.), mitochondrially related to *S. carolitertii* but included in a nuclear subclade with other individuals from its hydrological basin. On the other hand, a more according pattern with the one we found in our analysis is expected if a mitochondrial introgression from southern lineage of *S. pyrenaicus* into northern lineage would occur.

The use of several nuclear genes has demonstrated to be accurate for obtaining consistent phylogenetic relationships in the Iberian species of the genus *Squalius* relative to studies using only one nuclear gene, as was demonstrated in previous studies (Waap *et al.*, 2010). Nevertheless, the concatenated phylogeny of the six nuclear genes analysed in this study could not incorporate gene tree conflicts, such as different coalescence processes, to solve the nuclear relationships involving *S. catellanus*, *S. carolitertii* and the northern lineage of *S. pyrenaicus*. Species tree analysis has been developed as an effective tool for solving discordances between gene trees due to these methods accommodate lineage sorting and uncertainties in the gene tree (Knowles and Carstens, 2007; O'Meara, 2010), as could be one of the primary source of discordance between mitochondrial and nuclear topologies in the Iberian *Squalius* analysed. However Bayesian species delimitation may be sensitive to ancestral population size priors, which may influence the accuracy of obtaining high posterior probability values and it (Zhang *et al.*, 2011). Besides, some of these methodologies, as the one carried out here, are not able to incorporate processes such as hybridization or genetic introgression in the analysis (Yu *et al.*, 2012). For this reason, although Bayesian analysis under the species coalescent model may be useful to assess hypotheses of species delimitation, information provided by concatenated analyses, especially when mitochondrial and nuclear information is considered, must not be discarded, especially when evolutionary reticulate processes are involved. The low posterior probability values found for relationships among the species *S. pyrenaicus*, *S. valentinus* and *S. malacitanus* in the species tree probably are caused by the discordance between mitochondrial and nuclear genes found in the northern lineage of *S. pyrenaicus*.

2. Phylogeography, genetic structure and diversity

With regards to phylogeographical structure, network analyses demonstrated the complex genetic structure of the species *Squalius pyrenaicus* relative to *S. carolitertii* and other Iberian *Squalius* species (Fig. 6), as had been previously stated (Sousa-Santos *et al.*, 2007). This difference is more evident considering only these two species, which indicates the different

evolutionary history between the lineages of both species, probably as consequence of different intrinsic species traits or because both lineages occur in regions with different geological history, factors that have proven to influence genetic structure and diversity (Sousa-Santos *et al.*, 2016). Intrinsic factors of *S. carolitertii* and *S. pyrenaicus* do not differ significantly, as both species present similar life history traits (Fernández-Delgado and Herrera, 1995; Maia *et al.*, 2006). Consequently, other historical factors have to be invoked to explain the different genetic structure pattern found in phylogeographical networks of both species.

Mitochondrial haplotype networks were congruent with phylogenetic topologies. Within *S. pyrenaicus*, the three phylogenetic lineages were identified as different phylogroups in the mitochondrial network, being the Sado Basin the most divergent, as demonstrated the high number of mutational steps as well. The two phylogenetic lineages of *S. carolitertii* were also identified as divergent phylogroups in mitochondrial network. Nevertheless, haplotype networks demonstrated the presence of mitochondrial intraspecific relationships that were not observed in phylogenetic trees due to methodological premises of both approaches and the sharing of haplotypes among different basins, such is the case of Guadiana-Gilao, Guadalhorce-Velez and Mondego-Alcoa basins. Indeed, haplotype networks have the advantage over phylogenetic trees in representing intraspecific relationships better than phylogenetic trees (Posada and Crandall, 2001). Mitochondrial network also showed a geographical subdivision within two of the larger basins of the Iberian Peninsula, Tagus and Guadalquivir, and the close genetic relationship of different stretches of these two basins with other Iberian hydrological basins such as the Lower Guadalquivir with Guadalete and Odiel basins or Upper Guadalquivir with Guadalfeo Basin, or the mitochondrial connection of the Upper Tagus and Ebro Basin, demonstrating therefore a strong geographical structure within the species *S. pyrenaicus*. This network also showed the lack of geographical structure in the northern lineage of the species *S. carolitertii*. AMOVA analyses and *N_m* values corroborated mitochondrial haplotype network results, revealing significant levels of genetic structure in the species *S. pyrenaicus* among and within basins, on one side, and that significant genetic structure in *S. carolitertii* only may be attributed to the presence of two divergent phylogenetic lineages but when genetic structure is assessed in each of these lineages independently the signal of genetic structure is erased.

On the other hand, although nuclear markers has been useful to uncover phylogeographical signal (Hare, 2001), even in freshwater fishes (Domínguez-Domínguez *et al.*, 2011; Perea and Doadrio, 2015) nuclear network based on one small nuclear fragment only was able to recovered three large haplogropus (Fig. 8), although sharing of alleles between *S. carolitertii* and *S. pyrenaicus* occurred in one of the three lineages, probably as consequence of incomplete lineage sorting (Hare *et al.*, 2001). Nevertheless, although more specific relationships

could not be solved in the nuclear phylogenetic network, it highlights the closer relationship of the northern lineage of *S. pyrenaicus* with *S. carolitertii* than with the southern lineage of *S. pyrenaicus*.

With regards to genetic diversity, patterns were also different between *S. carolitertii* and *S. pyrenaicus*, being these values higher in *S. pyrenaicus* than in *S. carolitertii*, which made extendable findings found for Portuguese populations of both species (Sousa-Santos *et al.*, 2016). Within the genus *Squalius*, the drainage area has been significantly correlated to the genetic diversity indices (Almada and Sousa-Santos, 2010), although this assumption is not extensible to the entire cyprinid family (Sousa-Santos *et al.*, 2016). Nevertheless, in our study this finding only was true for *S. pyrenaicus*, in which the highest values were found in the largest basins (Tagus, Guadiana and Guadalquivir) relative to the smallest ones, but not for *S. carolitertii*, in which the Mondego Basin showed relatively higher values of genetic diversity than in the larger Douro Basin.

In southern half of the Iberian Peninsula the effect of Mediterranean seasonality is stronger than in northern basins (Alvarez-Coberas *et al.*, 2005). Summer droughts are responsible for fragmentation of populations and alteration of demographic parameters as consequence of annual bottleneck due to the desiccation of large river stretches in these southern basins of the Iberian Peninsula, influenced by abiotic and biotic processes (Prenda and Gallardo, 1996; Magalhaes *et al.*, 2002; Magoulick and Kobza, 2003; Mathews and Marsh-Matthews, 2003; Dekar and Magoulick, 2007), thus favouring genetic drift to act and finally promoting genetic structure and higher genetic diversity than in northern Iberian populations, less affected by Mediterranean seasonality, as we found in our study in *S. pyrenaicus* in comparison to *S. carolitertii* populations. Significantly higher diversity values have been also found in southern latitudes with warmer conditions and Mediterranean-type rivers for the genus *Squalius* populations from Guadiana basins (Almada and Sousa-Santos, 2010), and it has been considered as driving factor for metapopulation systems for other Iberian cyprinid species, such is the case of *Anaecypris*, which shares distribution range with *S. pyrenaicus* (Sousa-Santos *et al.*, 2014). However, some studies fail to find significant effect of hydrological regime in genetic diversity for cyprinids studied as a group (Sousa-Santos *et al.*, 2016), which suggest that this pattern is species-dependent according to other features of species, due to this authors do find a higher percentage of private haplotypes in small-sized and non-migratory species, such is the case of *Anaecypris* or *S. pyrenaicus*. Therefore, the characteristics of Mediterranean climate and its unpredictability in the duration of the drought-flood cycles have favoured the maintenance of this high genetic structure and diversity in *S. pyrenaicus* populations.

Other Pleistocene climatic factors may be invoked as responsible for the low level of genetic diversity in the northern lineage of *S. carolितertii*. Although glaciation not acted as intense as do in central or northern Europe, in the Mediterranean peninsulas the effects of glacial cycles upon biological populations have been demonstrated, and the local extinction of populations has been described (Hewitt, 2004; Schmitt and Hewitt, 2004; Schmitt, 2007). Therefore, reduction of levels of genetic diversity in the northern lineage of *S. carolितertii* seems to be caused by northern depletion of populations as consequence of Pleistocene glaciations (Sousa-Santos *et al.*, 2007; Almada and Sousa-Santos, 2010). Bottlenecks occurred in this northern lineage of *S. carolितertii* probably was stronger than in cyclical population reduction in southern latitudes of the Iberian Peninsula caused by summer droughts, favouring a higher effect of genetic drift in these northern populations with the consequent loss of genetic diversity (Nei *et al.*, 1975). Therefore, strong bottlenecks occurred in this regions could led to the loss of genetic diversity as consequence of loss of genotypes (Hewitt, 2004). Network haplotype structure suggests the presence of only one glacial refuge in this northern lineage, from which populations expand their range following paleogeographical configuration of the hydrological basins when climatic conditions became warmer. This issue implies that the lack of a clear genetic structure in the northern genetic lineage of *S. carolितertii* may be associated to the influence of Pleistocene glaciations, as the tMRCA for this northern lineage is lesser than 2.5 Ma. Similar lower levels of northern populations relative to southern populations in the Iberian Peninsula associated to Pleistocene glaciations have been reported for other vertebrate species (Alexandrino *et al.*, 2000; Gutiérrez-Rodríguez *et al.*, 2017). Nonetheless, higher levels of genetic values found in the Mondego-Alcoa lineage in *S. carolितertii* could be explained by the same reasons of influence of Mediterranean seasonality as explained for *S. pyrenaicus*.

3. Biogeographical patterns within *S. carolितertii* and *S. pyrenaicus*

Due to the mitonuclear discordance found in *S. pyrenaicus* relationships, historical biogeographical patterns of speciation processes were based on nuclear relationships, as this marker accurately reflected the evolutionary history of Iberian *Squalius* species. Nevertheless, we also estimated the age of cladogenetic events in the *MT-CYB* gene obtained from the mitochondrial plus nuclear concatenated species tree in order to confer a time framework to the inferred mitochondrial introgression occurred between *S. pyrenaicus* lineages, as well as to estimate the biogeographical patterns that showed the different lineages identified in the species *S. carolितertii* and *S. pyrenaicus* and their populations.

After the first evident divergence of the Portuguese *S. aradensis* / *S. torgalensis* lineage in the Lower Miocene period, ancient diversification of the main different *Squalius* lineages is not

easily understandable, as statistical support for these basal relationships was low in the nuclear species tree analysis ($pp=0.81$). This result suggests a reticulate pattern of diversification of these main Iberian lineages associated to complex tectonic movements from orogenic episodes (De Vicente *et al.*, 2011; Torné *et al.*, 2015), or that probably extinction events, difficult to estimate due to the incomplete fossil record, mask the real patterns of diversification of these old lineages, taking into account the long branch lengths observed in tree topology (Hey, 1992). Nevertheless, although nuclear relationships were not clearly solved three well supported lineages could be identified, which means that accurate historical biogeographic patterns may be explained.

From north to south, one of this lineage included *S. carolitertii* (northern and Mondego lineages) and northern lineage of *S. pyrenaicus*. Uplift of the proto-central systems started the differentiation of the Douro Basin (Portero and Aznar, 1984). During the Eocene, a northwest-southeast compression increased the uplift and generated the central system, which started to separate Douro from Tagus Cenozoic basins (Madrid Basin). During the late Eocene and Oligocene rapid uplift of the northern and northeast margins of the Douro Basin and the complete establishment of the Central System occurred, which as consequence of rapid erosion of uplifted areas refill Douro basin with sediments (Mediavilla *et al.*, 1994). Contemporary to this period, the Mondego Basin also constituted a Tertiary sub-basin isolated from the Lower Tagus sub-basin (Lisbon Basin) (Cunha, 1992). Nevertheless, these are very old paleogeographical events to explain the divergence of Douro and Tagus lineages and more recent paleogeographical events must be responsible to explain Douro-Tagus vicariance. From these old periods on, Pre-Douro was capturing small endorheic basins westward until acquired its final configuration in Middle-Upper Miocene (Armenteros-Armenteros, 1986; Pereira *et al.*, 2000). The Tagus Basin reaches its current configuration during the Upper Miocene (11-5 Ma) through the connection of the Loranca, Madrid and Lisbon basin systems as consequence of the tectonic movements that arise and posteriorly bent/inclined the Iberian Peninsula towards the Atlantic Ocean (López-Martínez *et al.*, 1989; Filipe *et al.*, 2009; de Vicente *et al.*, 2011) and probably was this time when *S. pyrenaicus* (including Zezere population) and *S. castellanus* diverged from *S. carolitertii*. Indeed, although change from exorheism to endorheism in the majority of Iberian basins took place in the Middle Tortonian, approximately 9.5 Ma ago, the end of exorheic processes occurred in Plio-Pleistocene (De Vicente *et al.*, 2011). According to that, the isolation of the Mondego basin has been related to the development of the current hydrological basins in Plio-Pleistocene when river terraces were formed as consequence of glacio-eustasy events (Ramos *et al.*, 2012). Sea level global changes were also responsible for the formation of fluvial terraces in coastal areas of Portugal during Quaternary period involving among other rivers Minho, Limia, Cavado and Ave drainages, which ended their final configuration at this time (Pereira *et al.*, 2000).

Pleistocene expansions from Tagus to Ebro and to small basins close to the Tagus mouth have been inferred. Headwaters of Tagus and Ebro River are very close and contacts between these two basins during hydrological dynamic alteration caused by glaciations are plausible (Durand *et al.*, 1999). On the other hand, continental shelf in the mouth of Tagus River is wide and shallow and downstream connections between Tagus River and these small basins were also possible during sea-level regression associated to Pleistocene glaciations (Alley *et al.*, 2005; Miller *et al.*, 2005), as occurred in other regions of Iberian Peninsula for other species of the genus *Squalius* (Perea and Doadrio, 2015). Later isolation of these small basins close to the Tagus mouth from the main Tagus basin favoured their isolation and the appearance of private mitochondrial haplotypes. Similar river capture may be invoked in the population of *S. carolitertii* from the headwaters of the Alberche River (Tagus Basin).

Within the *S. pyrenaicus* populations from the southern half of Iberian Peninsula, a first vicariance between the Guadiana and Gilão basins from the rest of the basins was inferred to take place in the Upper Miocene – Lower Pliocene. During Miocene the current drainage of the Guadiana Basin formed a lacustrine endorheic paleodrainage that at started to change to exoreic conditions as alpine tectonic activity caused the reactivation of a NE-SW fault system, which started to shaped an incise valley towards the Atlantic Ocean, (Moya-Palomares *et al.*, 2000; Garzón, 2005; Cunha *et al.*, 2012; Pais *et al.*, 2012). At the same time, the Guadalquivir Basin was created as a foreland basin following the final uplifting of the Betic Mountains in the Upper Miocene period, these tectonic movements resulted in the emergence of island and seaways that were reconfigured until end up in the current configuration of the Guadalquivir Basin (Braga *et al.*, 2003; Martin *et al.*, 2009, 2014). These tectonic movements in the southern half of the Iberian Peninsula associated to advanced stages of the Alpine orogeny probably were responsible for the divergence of the Guadiana – Gilão and Guadalquivir lineages. Later during Pliocene, an intrabasin differentiation explains the presence of a genetic substructure within the Guadalquivir Basin, with a segregated haplogroup from the lower reaches of the Guadalquivir Basin (Cala R.), related to Guadalete and Odiel basins, and other haplogroup from the upper reaches of the basin related to Guadalhorce – Velez, Segura and Guadalfeo basins, as showed haplotype network reconstructed. Connections between Upper Guadalquivir and Segura basins have been demonstrated for other cyprinids from the south of the Iberian Peninsula, such is the case of the barbel species *L. sclateri* (Casal-López *et al.*, in prep.).

Our biogeographical ancestral reconstruction inferred several colonization events during Upper Pliocene – Lower Pleistocene from the Guadalquivir to other southern Iberian basins according to this geographical substructure found. These results are congruent with previous biogeographic studies considering southern *Squalius* species (Perea *et al.*, 2016). Contacts

among headwaters of Guadalquivir tributaries and different southern basins was the probably explanation for these dispersal events. Headwaters of some of these rivers are very close or are including within the same geological region, as occurs with Guadalquivir and Segura rivers (Moral *et al.*, 2008), favouring contact among different hydrological basins. Besides, evidence of stream piracy is common in eastern areas of the Betic Mountains (Wenzens and Wenzens, 1995; Calvache and Viseras, 1997). Later vicariance between all this southern basins as the current southern Iberian hydrological basins were configured promoted the appearance of mitochondrial private haplotypes in all these southern basins.

CONCLUSIONS

This study highlights the importance of using multilocus mitochondrial and nuclear and species tree approaches to unravel phylogenetic relationships and hypothetical hybridization events between different species, as we inferred for the species *S. pyrenaicus*, which appeared to be polyphyletic due to the presence of three divergent lineages. Thus, further specific taxonomic studies are vital to decipher Iberian *Squalius* phylogeny. We also found a strong link between Cenozoic Iberian paleogeography and diversification of the different lineages belonging to *S. carolitertii* and *S. pyrenaicus*. On the other hand, historical climatic factors exerted a strong influence in genetic structure and diversity patterns in both species of *Squalius* considered in this study. Therefore, Mediterranean seasonality favoured a metapopulation structure and high levels of genetic diversity in *S. pyrenaicus*, whereas Pleistocene glaciations are responsible for low levels of genetic diversity found in *S. carolitertii*, a only glacial refuge is identify in this species, as is supported by the lack of genetic structure in its northern lineage and the estimate of its MRCA in 2.5 Ma and by the dispersal events from Douro to other basins from the northwestern quadrant of Iberian Peninsula.

REFERENCES

- Alexandrino J, Froufe E, Arntzen JW, Ferrand N. 2000. Genetic subdivision, glacial refugia and postglacial recolonization in the golden-striped salamander, *Chioglossa lusitanica* (Amphibia: Urodela). *Molecular Ecology*, 9: 771-781.
- Almada V, Sousa-Santos C (2010) Comparisons of the genetic structure of *Squalius* populations (Teleostei, Cyprinidae) from rivers with contrasting histories, drainage areas and climatic conditions based on two molecular markers. *Molecular Phylogenetics and Evolution*, 57, 924-931.
- Armenteros Armenteros I. 1986. Evolución paleogeográfica miocena en el sureste de la depresión del Ebro. *Studia Geologica Salmanticensia*, XXIII: 325-337.
- Braga, J.C., Martín, J.M., Quesada, C., 2003. Patterns and average rates of late Neogene-Recent uplift of the Betic Cordillera, SE Spain. *Geomorphology*, 50, 3-26.
- Calvo, J.P., Daams, R., Morales, J., López-Martínez, N., Augustini, J., Anadon, P., Armenteros, I., Cabrera, L., Civis,

- J., Corrochano, A., Díaz-Molina, M., Muñoz, A., Pérez-García, A., Pérez-González, A., Portero, J.M., Santisteban, C., Torres, T., Van der Meulen, A.J., Vera, J.A., Mein, P., 1993. Up-to-date Spanish continental Neogene synthesis and paleoclimatic interpretation. *Revista de la Sociedad Geológica de España*, 6, 29-40.
- Cabral, J. & Ribeiro, A. (1990). Neotectonic studies in Portugal – the neotectonic map. *Bull. INQUA Neotectonics Comm.*, 13, 6-8.
- Cabrera L, Gaudant J. 1985. Los Ciprínidos (Pisces) del sistema lacustre Oligocénico-Miocénico de los Monegros (secto SE de la cuenca del Ebro, provincias de Lleida, Tarragona, Huesca y Zaragoza). *Acta Geológica Hispánica*, 1985; 20(3-4): 219-226
- Carmona, J.A., Domínguez, J., Doadrio, I., 2000. Congruence between allozyme and cytochrome b gene sequence data in assessing genetic differentiation within the Iberian endemic *Chondrostoma lemmingii* (Pisces: Cyprinidae). *Heredity* 84, 721–732.
- Cloetingh, S., Burov, E., Beekman, F., Andeweg, B., Andriessen, P.A.M., García-Castellanos, D., De Vicente, G. & Vegas, R. (2002) Lithospheric folding in Iberia. *Tectonics*, 21(5), 1041-1067.
- Coelho MM, Bogutskaya NG, Rodrigues JA, Collares-Pereira MJ. 1998. *Leuciscus torgalensis* and *Leuciscus aradensis*, two new cyprinids from Portuguese fresh waters. *Journal of Fish Biology*, 52(5): 937-950.
- Cunha PP. 1992. *Estratigrafia e Sedimentologia dos Depósitos do Cretácico Superior e Terciário de Portugal Central, a Leste de Coimbra*. PhD Thesis. Faculdade Ciências Tecnologia. Universidade Coimbra. Coimbra.
- Cunha, C., Coelho, M.M., Carmona, J.A. & Doadrio, I. 2004. Phylogeographical insights into the origins of the *Squalius alburnoides* complex via multiple hybridization events. *Molecular Ecology*, 13(9): 2807-2817.
- Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* 9(8): 772.
- De Bruijine, C.H. & Andriessen, P. (2002) Far field effects of alpine plate tectonicism in the Iberian microplate recorded by fault-related denudation in the Spanish Central System. *Tectonophysics*, 349, 161-184.
- De Vicente, G., Cloetingh, S., Van Wees, J.D. & Cunha, P.P. (2011) Tectonic classification of Cenozoic Iberian forelands basins. *Tectonophysics*, 502, 38-61. Ferreira-Soares, A., Fonseca-Marques, J., Rocha, R.E.B., Cunha, P.P., Pinto-Duarte, L.V., Sequeira, A.D.J., Bernardo de Sousa, M., Gama-Pereira, L.C., Gomes, E., Pereira, E. & Rola dos Santos, J. (2005) Carta Geológica de Portugal, Folha 19-D, Coimbra-Lousã 1:50000, Instituto de Engenharia, Tecnologia e Inovação.
- Doadrio I. 1988. Delimitation of areas in the Iberian Peninsula on the basis of freshwater fishes. *Bonner zoologische Beiträge*, 39: 113-128.
- Doadrio, I. & Carmona, J. A. 2003. Testing freshwater Lago Mare dispersal theory on the phylogeny relationships of Iberian cyprinid genera *Chondrostoma* and *Squalius* (Cypriniformes, Cyprinidae). *Graellsia*, 59(2-3): 457-473.
- Doadrio I, Carmona JA (2004) Phylogenetic relationships and biogeography of the genus *Chondrostoma* inferred from mitochondrial DNA sequences. *Molecular Phylogenetics and Evolution*, 33, 802-815.
- Doadrio I, Carmona JA (2006) Phylogenetic overview of the genus *Squalius* (Actinopterygii, Cyprinidae) in the Iberian Peninsula, with description of two new species. *Cybium*, 30, 199-214.
- Doadrio, I., Perea, S., & Alonso, F. 2007. A new species of the genus *Squalius* Bonaparte, 1837 (Actinopterygii, Cyprinidae) from the Tagus River Basin (Central Spain). *Graellsia* 63(1): 89-100.
- Doadrio, I., Perea, S., Garzón-Heydt, P. & González, J.L. (2011) Ictiofauna Continental Española. Bases para su seguimiento. DG Medio Natural y Política Forestal. MARM. Madrid, 616 pp.
- Dupanloup, I., Schneider, S., Excoffier, L., 2002. A simulated annealing approach to define the genetic structure of populations. *Molecular Ecology*, 11: 2571-2581.
- Durand JD, Templeton AR, Guinand B, Imsiridou A, Bouvet Y. 1999. Nested clade and phylogeographic analyses of the chub, *Leuciscus cephalus* (Teleostei, Cyprinidae), in Greece: implications for Balkan Peninsula Biogeography. *Molecular Phylogenetics and Evolution*, 13(3): 566-580.
- Excoffier, L. & Lischer, H.E., 2010. Arlequin suite v.3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* 10(3), 564-567.

- Fernández-Delgado C, Herrera M. 1995. Age structure, growth and reproduction of *Leuciscus pyrenaicus* in an intermittent stream in the Guadalquivir river basin, Southern Spain. *Journal of Fish Biology*, 46: 371-380.
- Filipe AF, Araujo MB, Doadrio I, Angermeier PL, Collares-Pereira MJ. 2009. Biogeography of Iberian freshwater fishes revisited: the roles of historical versus contemporary constraints. *Journal of Biogeography*, 36: 2096-2110.
- Gaudant J. Contributions à la Paléontologie du Miocène moyen continental du bassin du Tage. II-Observations sur les dents pharyngiennes de Poissons Cyprinidés-Póvoa de Santarém. *Ciencias da Terra*, 1977; 3: 129-141.
- Gutiérrez-Rodríguez J, Barbosa AM, Martínez-Solano I. 2017. Present and past climatic effects on the current distribution and genetic diversity of the Iberian spadefoot toad (*Pelobates cultripes*): an integrative approach. *Journal of Biogeography*, 44(2): 245-258.
- Hare MP. 2001. Prospects for nuclear gene Phylogeography. *Trends in Ecology and Evolution*, 16(12): 700-706.
- Henriques R, Sousa V, Coelho MM. 2010. Migration patterns counteract seasonal isolation of *Squalius torgalensis*, a critically endangered freshwater fish inhabiting a typical Circum-Mediterranean small drainage. *Conservation Genetics*, 11(5): 1859–1870.
- Hewitt GM. 2004. Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London, Series B.*, 359: 183-195.
- Hey J. 1992. Using phylogenetic trees to study speciation and extinction. *Evolution*, 46(3): 627-640.
- Husemann M, Ray JW, King RS, Hooser EA, Danley PD. 2012. Comparative biogeography reveals differences in population genetic structure of five species of stream fishes. *Biological Journal of the Linnean Society*, 107: 867-885.
- Huson DH, Bryant D. 2006. Application of Phylogenetic Networks in Evolutionary Studies. *Mol. Biol. Evol.* 23(2), 254-267.
- Knowles LL, Carstens BC. 2007. Delimiting species without monophyletic gene trees. *Systematic Biology*, 56: 887-895.
- Landis MJ, Matzke NJ, Moore BR, et al., 2013. Bayesian analysis of biogeography when the number of areas is large. *Systematic Biology*, 62(6): 789-804.
- Lanfear, R., Calcott, B., Ho S.Y.W. & Guindon, S., 2012. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Mol. Biol. Evol.* 29(6), 1695-1701.
- Librado, P. & Rozas, J., 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25, 1451–1452. Doi: 10.1093/bioinformatics/btp187
- Lopes-Cunha, M., Aboim, M.A., Mesquita, N., Alves, M.J., Doadrio, I. & Coelho, M.M. 2012. Population genetic structure in the Iberian cyprinid fish *Iberochondrostoma lemmingii* (Steindachner, 1866): disentangling species fragmentation and colonization processes. *Biological Journal of the Linnean Society*, 105: 559–572.
- Maia HMS, Maia CFQ, Pires DFC, Valente ACN. 2006. Biology of the Iberian chub (*Squalius carolitertii*) in an atlantic-type stream (river Lima basin-north Portugal). A preliminary approach. *Limnetica*, 25(3): 713-722.
- Magalhães MF, Beja P, Canas C, Collares-Pereira MJ. 2002. Functional heterogeneity of dryseason refugia across a Mediterranean catchment: the role of habitat and predation. *Freshwater Biology*, 47: 1919-1934
- Martín, J.M., Braga, J.C., Aguirre, J., Puga-Bernabéu, A., 2009. History and evolution of the North-Betic Strait (Prebetic Zone, Betic Cordillera): a narrow, early Tortonian, tidal-dominated, Atlantic-Mediterranean marine passage. *Sediment. Geol.* 216, 80–90.
- Martín, J.M., Puga-Bernabéu, A., Aguirre, J., Braga, J.C., 2014. Miocene Atlantic-Mediterranean seaways in the Betic cordillera (Southern Spain). *Revista de la Sociedad Geológica de España* 27(1): 175-186.
- Mediavilla R, Martín-Serrano A, Dabrio CJ, Santisteban JI. 1994. En: Gierlowski-Kordesch E, Kelts K (eds.). Cenozoic lacustrine deposits in the Duero Basin (Spain). Global geological record of lake basins.
- Mesquita, N., Hänfling, B., Carvalho, G.R. & Coelho, M.M. 2005. Phylogeography of the cyprinid *Squalius aradensis* and implications for conservation of the endemic freshwater fauna of southern Portugal. *Molecular Ecology*, 14(7): 1939-1954.
- Mesquita, N., Cunha, C., Carvalho, G.R. & Coelho, M.M. 2007. Comparative phylogeography of endemic cyprinids in

- the south-west Iberian Peninsula: evidence for a new ichthyogeographic area. *Journal of Fish Biology*, 71: 45-75.
- Myers G.S. 1966. Derivation of freshwater fish from Central America. *Copeia*, 4, 766-773.
- Nei M, Maruyama T, Chakraborty R. 1975. The bottleneck effect and genetic variability in populations. *Evolution*, 29(1): 1-10.
- O'Meara BC. 2010. New heuristic method for joint species delimitation and species tree inference. *Systematic Biology*, 59: 59-73.
- Pais, J., Cunha, P.P., Pereira, D., Legoinha, P., Dias, R., Moura, D., Brum da Silveira, A., Kullberg, J.C., González-Delgado, J.A. (2012) The Paleogene and Neogene of Western Iberia (Portugal). A Cenozoic record in the European Atlantic domain. Springer Briefs in Earth Sciences. Springer, Berlin.
- Perea, S., Böhme, M., Zupancic, P., Freyhof, J., Sanda, R., Ozulug, M., Abdoli, A., Doadrio, I., 2010. Phylogenetic relationships and biogeographical patterns in Circum-Mediterranean subfamily Leuciscinae (Teleostei, Cyprinidae) inferred from both mitochondrial and nuclear data. *BMC Evol. Biol.* 10 (1), 265.
- Perea, S., Cobo-Simón, M. & Doadrio, I. 2016. Cenozoic tectonic and climatic events in southern Iberian Peninsula: Implications for the evolutionary history of freshwater fish of the genus *Squalius* (Actinopterygii, Cyprinidae). *Molecular Phylogenetics and Evolution*, 97: 155-169.
- Perea, S., Doadrio, I., 2015. Phylogeography, historical demography and habitat suitability modelling of freshwater fishes inhabiting seasonally fluctuating Mediterranean river systems: a case study using the Iberian cyprinid *Squalius valentinus*. *Mol. Ecol.* 24, 3706-3722.
- Pinheiro LM, Wilson RCL, Pena dos Reis R, Whitmarsh RB, Ribeiro A. 1996. The western Iberian margin: a geophysical and geological overview. In: Whitmarsh RB, Sawyer DS, Klaus A, Masson DG (eds.). *Proceedings of the Ocean Drilling Program. Scientific Results*, Vol. 149.
- Pereira DI, Alves MIC, Araújo MA, Cunha PP. 2000. Estratigrafia e interpretação paleogeográfica do Cenozóico continental do norte de Portugal. *Ciencias da Terra (UNL)*, 14: 73-84.
- Posada D, Crandall KA. 2001. Intraspecific gene genealogies: trees grafting into networks. *Trends in Ecology and Evolution*, 16(1): 37-45.
- Prenda J, Gallardo A. 1996. Self-purification, temporal variability and the macroinvertebrate community in small lowland Mediterranean streams receiving crude domestic sewage effluents. *Archiv für Hydrobiologie* 136:159–170.
- Quesada, C. (1991) Geological constraints on the Paleozoic tectonic evolution of tectonostratigraphic terranes in the Iberian massif. *Tectonophysics*, 185, 225-245.
- Rambaut, A. & Drummond, A.J., 2007. Tracer v1.4. Available from <http://beast.bio.ed.ac.uk/Tracer>.
- Ramos AM, Cunha PP, Cunha LS, Gomes A, Lopes FC, Buylaert J-P, Murray AS. 2012. The River Mondego terraces at the Figueroa da Foz coastal area (western central Portugal): geomorphological and sedimentological characterization of a terrace staircase affected by differential uplift and glacio-eustasy. *Geomorphology*, 165-166: 107-123.
- Rice, W.R., 1989. Analyzing tables of statistical tests. *Evolution*, 223-225.
- Robalo, J.I., Santos, C.S., Almada, V.C. & Doadrio, I. (2006) Paleobiogeography of two Iberian endemic cyprinid fishes (*Chondrostoma arcasii* – *Chondrostoma macrolepidotus*) inferred from mitochondrial DNA sequence data. *Journal of Heredity*, 97, 143–149.
- Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, L., Liu, L., Suchard, M.A. & Huelsenbeck, J.P., 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 61(3), 539:542. Doi: 10.1093/sysbio/sys029
- Sanjur, O.I., Carmona, J.A. & Doadrio, I. 2003. Evolutionary and biogeographical patterns within Iberian populations of the genus *Squalius* inferred from molecular data. *Molecular Phylogenetics and Evolution* 29, 20–30.
- Schmitt T, Hewitt G. 2004. The genetic pattern of populations threat and loss: a case study of butterflies. *Molecular Ecology*, 13: 21-31.
- Slatkin, M. & Barton, N.H., 1989. A comparison of three indirect methods for estimating average levels of gene flow.

- Evolution, 43(7), 1349–1368. Doi: 10.2307/2409452
- Sousa-Santos, C., Robalo, J.I., Pereira, A.M., Branco, P., Santos, J.M., Ferreira, M.T., Sousa, M., Doadrio, I. 2016. Broad-scale sampling of primary freshwater fish populations reveals the role of intrinsic traits, inter-basin connectivity, drainage area and latitude on shaping contemporary patterns of genetic diversity. *PeerJ*. DOI 10.7717/peerj.1694.
- Stamakis A. 2006. RaxML-VI-HPC: Maximum Likelihood-based phylogenetic Analyses with thousands of taxa and Mixed Models. *Bioinformatics*, 22(21), 2688–2690.
- Stamakis A, Blagojevic F, Nikolopoulos D, Antonopoulos C., 2007. Exploring new search algorithms and hardware for phylogenetics: RaxML Meets the IBM Cell. *The Journal of VLSI Signal Processing*, 48, 271-286.
- Swofford DL. 2003. PAUP*. Phylogenetic Analysis Using Parsimony (*and other methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.
- Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22(22), 4673–4680.
- Vergés, J., Fernández, M. (2006) Ranges and basins in the Iberian Peninsula: their contribution to the present topography. *European lithosphere dynamics: Geological Society, London. Memoirs* (ed. by Gee, G.D. & Stephenson, R.A.). pp 223-234.
- Zardoya, R., Doadrio, I. 1998. Phylogenetic relationships of Iberian cyprinids: systematic and biogeographical implications. *Proceedings of the Royal Society of London B*, 265: 1365-1372.
- Zhang C, Zhang D-X, Zhu T, Yang Z. 2011. Evaluation of a Bayesian coalescent method of species delimitation. *Systematic Biology*, 60(6): 77-761.

SUPPLEMENTARY MATERIAL

Tabla S1. Sampling specimens, localities and GenBank accession numbers. Gene abbreviations: *MT-CYB*, mitochondrial cytochrome *b*, *RAG1*: nuclear recombination activating gene 1; *S7*, ribosomal protein S7; *EFA1α*, elongation factor 1 alpha; *RHO*, rhodopsin gene; *ACTB*, actin beta gene; *EGR2b*, early growth response protein gene 2b.

River / Basin	Sample size (<i>MT-CYB</i> / <i>RAG1</i> / <i>S7</i> / <i>EFA1α</i> / <i>RHO</i> / <i>ACTB</i> / <i>EGR2b</i>)	GenBank accession numbers (pendientes de asignación)
<u>Squalius pyrenaicus</u>		
Baixas / Ebro		xxx-xxx
Piedra / Ebro		xxx-xxx
Jerea / Ebro		xxx-xxx
Lis / Lis		xxx-xxx
Lizandro / Lizandro		xxx-xxx
Samarra / Samarra		xxx-xxx
Colares / Colares		xxx-xxx
Lage / Lage		xxx-xxx
Jamor / Jamor		xxx-xxx
Acebo / Tagus		xxx-xxx
Alburrel / Tagus		xxx-xxx
Alagón / Tagus		xxx-xxx
Almonda / Tagus		xxx-xxx
Alviela / Tagus		xxx-xxx
Arrago / Tagus		xxx-xxx
Aurela / Tagus		xxx-xxx
Batuecas / Tagus		xxx-xxx
Cedena / Tagus		xxx-xxx
Cofio / Tagus		xxx-xxx
Erjas / Tagus		xxx-xxx
Escabas / Tagus		xxx-xxx
Gévalo / Tagus		xxx-xxx
Grande da Pipa / Tagus		xxx-xxx
Guadiela / Tagus		xxx-xxx
Gualija / Tagus		xxx-xxx
Hurdano / Tagus		xxx-xxx
Huso / Tagus		xxx-xxx
Ibor / Tagus		xxx-xxx
Jarama / Tagus		xxx-xxx
Jerte / Tagus		xxx-xxx
Maior / Tagus		xxx-xxx
Muge / Tagus		xxx-xxx
Nabão / Tagus		xxx-xxx
Ocreza / Tagus		xxx-xxx
Ompolveda / Tagus		xxx-xxx

Ota / Tagus	XXX-XXX
Pesquero / Tagus	XXX-XXX
Ponsul / Tagus	XXX-XXX
Ramacastañas / Tagus	XXX-XXX
Rivera de Bronco	XXX-XXX
Rivera de Gata	XXX-XXX
San Blas (arroyo) / Tagus	XXX-XXX
Sanmartin / Tagus	XXX-XXX
Sever / Tagus	XXX-XXX
Tajuña / Tagus	XXX-XXX
Tietar / Tagus	XXX-XXX
Trevejana / Tagus	XXX-XXX
Vid / Tagus	XXX-XXX
Zezere / Tagus	XXX-XXX
Campilhas / Sado	XXX-XXX
Odivelas / Sado	XXX-XXX
São Martinho / Sado	XXX-XXX
Arronches / Guadiana	XXX-XXX
Almargem / Guadiana	XXX-XXX
Ardila / Guadiana	XXX-XXX
Azuer / Guadiana	XXX-XXX
Caia / Guadiana	XXX-XXX
Calaboza / Guadiana	XXX-XXX
Cheleiros / Guadiana	XXX-XXX
Cigüela / Guadiana	XXX-XXX
Cobres / Guadiana	XXX-XXX
Esteras / Guadiana	XXX-XXX
Gévora / Guadiana	XXX-XXX
Guadalemar / Guadiana	XXX-XXX
Guadalmez / Guadiana	XXX-XXX
Guadarranque / Guadiana	XXX-XXX
Matachel / Guadiana	XXX-XXX
Murtigas / Guadiana	XXX-XXX
Odeleite / Guadiana	XXX-XXX
Oeiras / Guadiana	XXX-XXX
Quejigares / Guadiana	XXX-XXX
Ruecas / Guadiana	XXX-XXX
Ruidera (lagunas) / Guadiana	XXX-XXX
Sillo / Guadiana	XXX-XXX
Siruela / Guadiana	XXX-XXX
Valdeazogues / Guadiana	XXX-XXX
Valparairo / Guadiana	XXX-XXX
Vascão / Guadiana	XXX-XXX
Zújar / Guadiana	XXX-XXX
Guadiana / Guadiana	XXX-XXX
Gilão	XXX-XXX

Villar / Odiel	XXX-XXX
Ciudadreja / Guadalquivir	XXX-XXX
Cabrera / Guadalquivir	XXX-XXX
Cala / Guadalquivir	XXX-XXX
Fresneda / Guadalquivir	XXX-XXX
Guadalmena / Guadalquivir	XXX-XXX
Manzano / Guadalquivir	XXX-XXX
Molinos / Guadalquivir	XXX-XXX
Montemayor / Guadalquivir	XXX-XXX
Montoro / Guadalquivir	XXX-XXX
Retortillo / Guadalquivir	XXX-XXX
Rivera de Huelva / Guadalquivir	XXX-XXX
Rivera de Hueznar / Guadalquivir	XXX-XXX
Robledillo / Guadalquivir	XXX-XXX
Viar / Guadalquivir	XXX-XXX
Yeguas / Guadalquivir	XXX-XXX
Segura / Segura	XXX-XXX
Afluente del Guadalete / Guadalete	XXX-XXX
Guadalete / Guadalete	XXX-XXX
Grande / Guadalhorce	XXX-XXX
Pereilas / Guadalhorce	XXX-XXX
Turón / Guadalhorce	XXX-XXX
Cuevas / Velez	XXX-XXX
Sabar / Velez	XXX-XXX
Trévez / Guadalfeo	XXX-XXX
Guadalfeo / Guadalfeo	XXX-XXX
<u>Squalius carolitertii</u>	
Alcoa / Alcoa	XXX-XXX
Alva / Mondego	XXX-XXX
Arunca / Mondego	XXX-XXX
Ceira / Mondego	XXX-XXX
Corvo / Mondego	XXX-XXX
Dão / Mondego	XXX-XXX
Mortagua / Mondego	XXX-XXX
Águeda / Vouga	XXX-XXX
Caima / Vouga	XXX-XXX
Sul / Vouga	XXX-XXX
Ave / Ave	XXX-XXX
Homen / Cavado	XXX-XXX
Neiva / Neiva	XXX-XXX
Pego / Pego	XXX-XXX
Salas / Limia	XXX-XXX
Limia / Limia	XXX-XXX
Otaiven / Otaiven	XXX-XXX
Lerez / Lerez	XXX-XXX
Arnoia / Minho	XXX-XXX

Mouro / Minho	xxx-xxx
Tea / Minho	xxx-xxx
Arlanza / Douro	xxx-xxx
Cea / Douro	xxx-xxx
Coa / Douro	xxx-xxx
Corgo / Douro	xxx-xxx
Curueño / Douro	xxx-xxx
Ebrillos / Douro	xxx-xxx
Eresma / Douro	xxx-xxx
Manzanas / Douro	xxx-xxx
Paiva / Douro	xxx-xxx
Razón / Douro	xxx-xxx
Sabor / Douro	xxx-xxx
Sousa / Douro	xxx-xxx
Támega / Douro	xxx-xxx
Távora / Douro	xxx-xxx
Torío / Douro	xxx-xxx
Turones / Douro	xxx-xxx
Valdavia / Douro	xxx-xxx
Valderaduey / Douro	xxx-xxx
Yeltes / Douro	
Other <i>Squalius</i> species:	
<i>S. aradensis</i>	xxx-xxx
<i>S. castellanus</i>	xxx-xxx
<i>S. malacitanus</i>	xxx-xxx
<i>S. torgalensis</i>	xxx-xxx
<i>S. valentinus</i>	xxx-xxx

The Dinaric karst is driving the diversification of the Eastern Adriatic *Squalius* Bonaparte, 1837 (Actinopterygii, Cyprinidae) species belonging to the Mediterranean genetic lineage: evidence from mitochondrial and nuclear markers.

(manuscrito en preparación; short communication)

RESUMEN

La región de la vertiente oriental del Adriático presenta una fauna de peces continentales muy antigua y con un alto grado de endemismos. Además, esta región tiene una configuración hidrogeomorfológica muy particular que viene determinada por el karst Dinárico, que se caracteriza por un régimen hidrogeológico e hidrológico que depende principalmente de la interacción entre el agua subterránea y el agua de superficie. Esta característica influye enormemente en la estructura de las comunidades acuáticas que habitan esta área, entre ellas los peces de agua dulce, los cuales pasan parte del tiempo bajo la superficie, dando lugar a complejos patrones evolutivos. El principal objetivo de este estudio es reconstruir las relaciones filogenéticas a nivel inter e intraespecífico de las especies adriáticas del género *Squalius* que pertenecen al linaje Mediterráneo a partir de datos mitocondriales y nucleares. Nuestros resultados soportan un escenario de procesos de especiación alopátrica reciente para estas especies de *Squalius* de acuerdo a la finalización de la formación del karst Dinárico en el Plioceno Superior – Pleistoceno Inferior. Las recientes estimas de divergencia para la separación de estas especies que están estadísticamente soportadas desde el punto de vista mitocondrial y nuclear apoyan esta idea de especiación reciente y que la formación de esta unidad geomorfológica ha sido el principal factor responsable de estos procesos de especiación, como también ha sido sugerido para otros grupos taxonómicos. Por otro lado, la falta de estructura intraespecífica en las especies *S. illyricus*, *S. squalius* y *S. microlepis* pone de manifiesto complejas relaciones hidrogeográficas entre las cuencas que conforman el karst Dinárico.

Keywords: Adriatic *Squalius* species, Dinaric karst, recent speciation

ABSTRACT

The Eastern Adriatic region has an old freshwater fish fauna with a high level of freshwater endemisms. Besides, this region has a particular geomorphological configuration determined by the Dinaric Alpine karst, which is characterized by hydrogeological and hydrological regime of that depends mostly upon the interaction between groundwater and surface water. This characteristic strongly influences the structure of aquatic communities inhabiting this area, among them freshwater fishes, which spend part of the time underground, leading in complex evolutionary patterns. The main goal of this study is to reconstruct the phylogenetic relationships at inter- and intraspecific level of the species belonging to the Mediterranean *Squalius* lineage that inhabit in the Eastern Adriatic region based on mitochondrial and nuclear markers. We obtained highly congruent mitochondrial and nuclear topologies supporting the identity of the five species inhabiting the Eastern Adriatic slope. Our results support a scenario of recent vicariant speciation for these *Squalius* species following the arising of the current hydrological basins that define the Dinaric karst since the Plio-Pleistocene period. The young divergence time estimations for the separation between these species from the Eastern Adriatic slope as well as the strong statistical support for mitochondrial and nuclear relationships found in Bayesian topologies suggest that speciation process have been driven by the formation of this geomorphological unit, as has been suggested for other taxonomic groups. On the other hand, the lack of a strong phylogeographical structure at intraspecific level for *S. illyricus*, *S. svallize* and *S. microlepis*, highlights the complex hydrogeographical structure of Dinaric karst.

Keywords: Adriatic *Squalius* species, Dinaric karst, recent speciation

INTRODUCTION

The Eastern Adriatic region has an old freshwater fish fauna, which comprises high level of endemisms as same as occurs with other freshwater taxa (Banarescu, 2004; Sket *et al.*, 2004; Sanda *et al.*, 2008; Buj *et al.*, 2010; Freyhof and Brooks, 2011). Besides, this region has a particular geomorphological configuration determined by the Dinaric Alpine karst, which constitutes an arc from the north of Italy since the Bay Trieste in Italy to Drim River Basin in Albania with a surface of approximately 60,000 km² (Bonnaci, 2013). The hydrogeological and hydrological regime of all water phenomena in the Dinaric karst depends mostly upon the interaction between groundwater and surface water. The dominant flow of the groundwater contained in the Dinaric karst is towards the Adriatic Sea through rivers and many permanent and intermittent coastal and submarine springs (Bonnaci, 2013). Therefore, the Eastern Adriatic slope is formed by short rivers in surface, with seasonal fluctuating flow, but that have a complex underground geomorphology with numerous sinkholes, caves, lakes and streams, most of them showing continuous groundwater flow, and numerous springs that emerge to surface (Assad and Jordan, 1994; Padalancic *et al.*, 2012b).

This characteristic strongly influences the structure of aquatic communities inhabiting this area by several reasons. First, some of the Eastern Adriatic freshwater species, not necessarily strict troglodytes or cave-dwellers (e.g. Culver and Sket, 2000), spend part of the time underground, among them freshwater fishes where the riverbeds bear permanently water (Leiner and Popovic, 1984; Padalancic *et al.*, 2012b). Hence, these species would be able to avoid natural perturbations such as Mediterranean summer droughts occurring in superficial waters at these latitudes (Benito *et al.*, 2008). Second, karstic environments promote complex hydrological basins and, as consequence, rivers that apparently form independent basins in the surface may be connected underground even along distant regions (Bonacci *et al.*, 2013) so that relationships among phylogenetic lineages are not concordant with present-day hydrological regimes (Veronik *et al.*, 2004; Zaksek *et al.*, 2009). Actually, some studies have showed evidence for underground connections and freshwater fish migration based on molecular approaches, as occur with *Delminichthys adpersus* (Heckel, 1843) (Padalancic *et al.*, 2012a,b), which in turn evidence the utility of analyzing genetic structure of freshwater organisms inhabiting in such particular region to understand its evolutionary history and the hydrological functioning of the Dinaric karst. For these reasons, the phylogenetic and phylogeographic relationships of freshwater species occurring in the Adriatic karstic region will be strongly influenced by this kind of habitat.

Although it is an old geomorphological structure, the formation of the present-day geomorphology of the Dinaric karst started at the end of the Pliocene-beginning of the Pleistocene following neotectonic movements that promoted the development of karstic areas and culminated with current Eastern Adriatic hydrogeographic configuration (Rodic, 1981;

Mihevć, 2001; Pavelić and Belak, 2008). Karstification changes topography and hydrology of freshwater systems (Bosák, 2008) and may produce either habitat fragmentation or habitat connection imposing different constraints to surface freshwater fauna in relation to their history life traits (Padalancić *et al.*, 2012; Previsić *et al.*, 2014).

Some fish species belonging to the freshwater genus *Squalius* are distributed within the region delimited by the Dinaric karst (Kottelat and Freyhof, 2007). This genus is one of the groups of freshwater fish more diversified in the Mediterranean basin. It has approximately fifty species, most of them endemic from southern Mediterranean peninsulas (Sanjur *et al.*, 2003; Perea *et al.*, 2010; Kottelat and Freyhof, 2007; Doadrio *et al.*, 2011; Eschemeyer *et al.*, 2016) and comprises two evolutionary lineages, Mediterranean and Centroeuropean, isolated one from each other since Middle Miocene (Sanjur *et al.*, 2003; Doadrio and Carmona, 2003; Perea *et al.*, 2010). One of these *Squalius* lineages, the Mediterranean, is studied in this paper. This lineage is distributed throughout southern Europe, where is restricted to a few river basins from the Iberian Peninsula, Italy, Eastern Adriatic slope, Greece and Western Anatolia (Kottelat and Freyhof, 2007; Doadrio *et al.*, 2011; Eschemeyer *et al.*, 2016). Mediterranean *Squalius* species are small-medium size fishes inhabiting streams and rivers with typical Mediterranean climate typology of summer droughts (Benito *et al.*, 2008), and most of them are seriously threatened, such is the case of *S. tenellus* Heckel, 1843, as consequence of the increasing degradation of Mediterranean freshwater ecosystems (Freyhof and Kottelat, 2008; Freyhof and Brooks, 2011).

From an evolutionary point of view, phylogenetic relationships of *Squalius* species belonging to the Mediterranean lineage have been broadly analyzed since the last fifteen years, especially in the Iberian Peninsula and Greece (Doadrio and Carmona, 1998, 2003, 2004; Durand *et al.*, 2000; Sanjur *et al.*, 2003; Perea *et al.*, 2010). Nevertheless, the genetic diversity and structure at intraspecific level is not well known for all Mediterranean species, such is the case of Western Balkans, in particular those from the Eastern Adriatic slope. Until now, few papers have deal with evolutionary questions in Eastern Adriatic *Squalius* species, but these authors only included one or two individuals per species for phylogenetic purposes (Durand *et al.*, 2000; Perea *et al.*, 2010). However, none of those studies addressed phylogenetic or phylogeographic structure of Eastern Adriatic *Squalius* at intraspecific level. Eastern Adriatic *Squalius* species belonging to the Mediterranean lineage are distributed throughout a restricted area in Bosnia-Herzegovina and Croatia (Fig. 1). This group is constituted by five species (*S. illyricus* Heckel and Kner, 1857, *S. microlepis* Heckel 1843, *S. svallize* Heckel and Kner, 1857, *S. tenellus* Heckel 1843 and *S. zrmanjae* Karaman, 1928), all of them endemisms from this region and closely related to the Italian species. Phylogenetic relationships showed that Eastern Adriatic species were a sister group of the Italian species *S. lucumonis* Bianco, 1983 and clustered together with the Greek species, *S. keadicus* Stephanidis, 1971 and *S. ghigii* Gianferrari, 1927, and the Iberian *Squalius* forming the Mediterranean lineage of the genus (Perea *et al.*, 2010). In

general the distribution range of these five Eastern Adriatic species is restricted to a few river basins (Kottelat and Freyhof, 2007; Caleta *et al.*, 2015) and we hypothesized they have been probably involved in an allopatric process of speciation conditioned by the geomorphology of the Dinaric Alpine karst during the end of the Tertiary and the Quaternary periods.

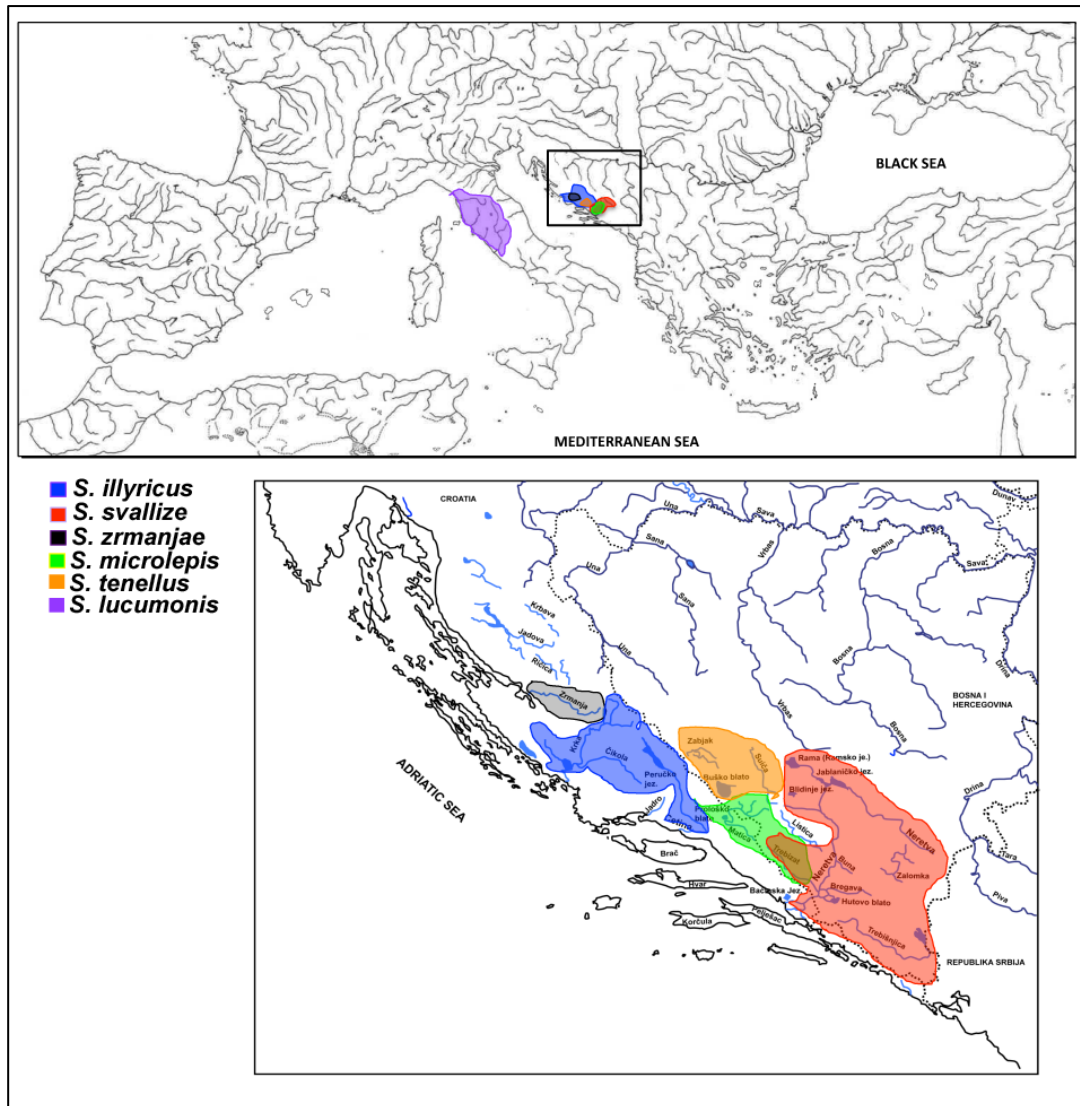


Fig. 1. Distribution range of Adriatic species of the genus *Squalius*.

Therefore, the main goal of this study is to reconstruct the phylogenetic relationships at inter- and intraspecific level of the species belonging to the Mediterranean *Squalius* lineage that inhabit in the Eastern Adriatic region based on mitochondrial and nuclear markers. Our hypothesis to test in this study is that taking into account the complexity of the hydrological basins of the Dinaric karst, the phylogeny and phylogeography of Eastern Adriatic species will be also complex and will be link to the history of this geomorphological formation, being allopatric speciation the main process acting in the diversification of these *Squalius* species. On the other hand, owing to the onset of the current configuration of this karstic region dates from Late

Pliocene-Early Pleistocene (Rodic, 1981; Pavelic and Belak, 2008) these period was probably crucial for the genetic diversification of Eastern Adriatic *Squalius* species. For this reason, a second objective of this study is to estimate the timing of diversification of these species and their populations in order to test the hypothesis of a rapid genetic diversification of these *Squalius* species following the Dinaric karst formation.

MATERIAL AND METHODS

1. Sampling

One hundred nineteen specimens of several *Squalius* species inhabiting Eastern Adriatic slope, along with *S. lucumonis* phylogenetically related to the former *Squalius* species, were collected with electrofishing under authorities permission (Fig. 1, Table 1). Other species belonging to the Mediterranean lineage were included in phylogenetic analysis. This represents an adequate sampling to the species: *Squalius illyricus*, *S. microlepis* and *S. svallize*. We failed to obtain samples of *Squalius zrmanjae*, which occurs in Krka and Zrmanja rivers (Freyhof and Kottelat, 2008; Caleta et al., 2015). In our field surveys we collected only some *Squalius* specimens in Krka River, but they were identified as *S. squalus* (not included in this study). For this reason samples of *S. zrmanjae* come from Genbank (Accession Number: HM560213 and AJ251093). To *S. tenellus* a species listed as Endangered (Freyhof and Kottelat, 2008) was found in low frequency in our field surveys in three localities. For this reason, as well as *S. zrmanjae*, they were only used for phylogenetic purposes but not for phylogeographical approaches. In order to obtain a better resolve tree topology we also included in the mitochondrial analyses other *Squalius* species as well as the genus *Petroleuciscus* as outgroup according to previous studies (Sanjur et al., 2003; Perea et al., 2010). Based in these studies, two species of the Euroasiatic lineage (*S. peloponensis* and *S. cephalus*) were used as outgroups.

2. DNA extraction, amplification and sequencing

Fin-clips were preserved in ethanol 95% immediately after DNA Collection of the National Museum of Natural Sciences of Madrid (MNCN-CSIC), Spain. For phylogenetic purposes we analyzed several individuals per population for mitochondrial *MT-CYB* (*MT-CYB*; 1140 bp). We also sequenced one individual per species for three nuclear genes (*RAG1*, *S7*, *EFA1*; total alignment: 2420 bp) for all species of the genus *Squalius* belonging to the Mediterranean lineage. Total genomic DNA was extracted from fin-clip tissues using the commercial kit Biosprint15 for tissue and blood (Qiagen). Primers and PCR protocols followed Perea et al., (2010) and Waap et al., (2010). After checking PCR products on 1% agarose gels, they were purified by ExoSAP-IT™ (USB) and directly sequenced on MACROGEN Inc (Amsterdam, The Netherlands;

<http://www.macrogen.com>) using a 3730XL DNA sequencer. All new sequences (haplotypes and alleles) generated in this paper were deposited on GenBank database (Accession Numbers: xxx-xxx).

Tabla 1. Localities sampled for *Squalius* species from Eastern Adriatic slope and *S. lucumonis*.

SPECIES	LOCALITY SAMPLED	NUMBER OF INDIVIDUALS
<i>S. illyricus</i>	Cetina R.	5
	Krka R.	8
	Peruca R.	2
	Cikotina R.	3
	Prancevici R.	2
<i>S. svallize</i>	Rama Lake	11
	Bilecko Lake	3
	Trebisnjica R.	4
	Neretva R.	1
	Jasenica R.	1
	Bregava R.	4
<i>S. tenellus</i>	Zabljac R.	2
<i>S. zrmanjae</i>	Zrmanja R.	2
<i>S. microlepis</i>	Tihaljina R.	8
	Prolosko Lake	6
	Ricina R.	2
	Ricice R.	3
	Trebizat R.	3
	Krenica Lake	4
	Grudsko polje	6
<i>S. lucumonis</i>	Tiber R.	2
	Arno R.	6

3. Phylogenetic and phylogeographic analyses

All sequences were initially aligned using the default pairwise and multiple alignment parameters in Clustal W (Thompson *et al.*, 1994) implemented in MEGA program v.6.0 (Tamura *et al.*, 2013). Chromatograms were verified and alignments were latterly revised by eye. Phylogenetic analyses were performed for mitochondrial and nuclear dataset separately and for concatenated dataset as well. For nuclear data and phylogeographical analyses, alleles were phased with PHASE algorithm implemented in DnaSP v5.0 (Librado and Rozas, 2009). Partitionfinder (Lanfear *et al.*, 2012) was used to find the best partition scheme and the best model of evolution for phylogenetic reconstruction.

Bayesian inference (BI) was performed using MrBayes v.3.2 (Ronquist *et al.*, 2012). Two simultaneous analyses were run along 10^7 generations, each with four MCMC chains sampling every 100 generations. Convergence was checked on Tracer v.1.6. (Rambaut and Drummond, 2007). After the first 10% of generations were discarded by burn-in, we obtained the 50% majority rule consensus tree and the posterior probabilities. Tree topologies were edited using FigTree v.1.4. software (<http://tree.bio.ed.ac.uk/software/figtree/>). On the other hand, we reconstructed haplotype network sequences to resolve relationships among closely related haplotypes (Crandall, 1994). Haplotype genealogies for the *MT-CYB* gene were obtained by HaploViewer software (available at <http://www.cibiv.at/~%20greg/haploviewer>). For phylogeographic approaches only *Squalius illyricus*, *S. microlepis* and *S. svallize* were used due to the small sample size of the species *S. tenellus* and *S. zrmanjae*.

4. Molecular clock

To estimate the main divergence events within the evolutionary history of Eastern Adriatic *Squalius* we carry out a molecular clock analysis in BEAST v.1.8.3. software (Drummond *et al.*, 2012). As a general clock-like behaviour was rejected ($p > 0.05$, analysis performed in MEGA v.6.0; Tamura *et al.*, 2013), divergence times and their credibility intervals at 95% (highest posterior density: HPD) were estimated using a relaxed clock model in which branch rates were drawn from an uncorrelated lognormal distribution (Drummond *et al.*, 2006). We estimate interspecific divergence times using *BEAST, which allows incorporating the coalescent process in the species tree estimation (Heled and Drummond, 2010). In order to incorporate two fossil calibrations to the analyses, we open the phylogeny to incorporate other *Squalius* species from which nuclear genes considered in this study are available (Perea *et al.*, 2010; Waap *et al.*, 2010). For calibrating the molecular clock, fossil evidence for *Squalius*: *Leuciscus antunesi* from Portugal dated from Middle Miocene (13.5-14.5 Ma; Cabrera and Gaudant, 1985) and *Leuciscus aff. cephalus* (6.5 Ma; Lava 2 deposit, Greece; Böhme and Ilg, 2003) was used to place the constraints of the age of different nodes within the interspecific topology. On the other hand, with the aim to estimate the tMRCA between populations we considered the *MT-CYB* gene tree obtained in the same *BEAST analysis. Two independent MCMC analyses were run 10^6 generations. Convergence was assessed using Tracer v.1.6 (Drummond and Rambaut, 2007) before to reconstruct the 50% majority-rule consensus tree of the post burn-in trees with the software TreeAnnotator v.1.8.3 (Drummond and Rambaut, 2007).

RESULTS AND DISCUSSION

In the world karstic areas occupy less than 1%, nevertheless, they have been showed to be relevant in evolutionary processes of organisms that inhabit them (Cohen, 2003). The Balkan Peninsula is well known for its variety of freshwater ecosystems being the Dinaric karst in the Eastern Adriatic slope one of its most significant regions for freshwater fauna. In fact, effects of karst development on hydrology in the region have been generally very drastic (Sket, 2002; Prelovsek, 2010).

Dinaric karst has been considered as a refuge for fauna leading in a high number of endemism and high level of diversity in freshwater species (Banarescu, 2004). Thus, high differentiation of mitochondrial haplogroups has been found in the Dinaric region for freshwater fishes as well as for some invertebrate species (Sket, 1999; Schletova *et al.*, 2004; Verovnik *et al.*, 2004, 2005; Trontelj *et al.*, 2005; Defavari *et al.*, 2012; Buj *et al.*, 2014). Nevertheless, not only a differentiation of Adriatic from European populations has occurred but also there has been described a small-scale local population differentiation within the Dinaric karst, which has been well documented for invertebrates such as mussels. Hence, speciation drove by Dinaric karst formation has been especially reported for mussels aquatic insects, in which local allopatry may be defined even at landscape scale (Schmidt-Kloiber and Hering, 2012; Bilandžija *et al.*, 2013; Klobucar *et al.*, 2013; Previsic *et al.*, 2014). Probably is also common in aquatic invertebrates associated to their small body size and low dispersal ability in relation to freshwater fishes (Klobucar *et al.*, 2013; Previsic *et al.*, 2014). Therefore, complex geographical features, as those found in the Dinaric karst, may harbor suitable microhabitats in the form of refugia. The survival of populations along evolutionary time in refuges allowed the accumulation of genetic differences, the development of genetic lineages and, ultimately, the promotion of speciation (Hewitt, 2004).

In the case of freshwater fishes, with a large size and greater mobility than invertebrates, genetic differentiation also may occur at local scale but more complex patterns are expected, such is the case of that showed by the *Squalius* species from the Eastern Adriatic slope belonging to the Mediterranean lineage, as we demonstrate here. Our results support a scenario of recent vicariant speciation for these *Squalius* species following the arising of the current hydrological basins that define the Dinaric karst since the Plio-Pleistocene period (Rodic, 1981; Pavelic and Belak, 2008). Therefore, we obtained highly congruent mitochondrial and nuclear topologies supporting the identity of the five species inhabiting the Eastern Adriatic slope (*S. illyricus*, *S. microlepis*, *S. svallize*, *S. tenellus* and *S. zrmanjae*) (Figs. 1 and 2). The main divergence times estimated by our molecular clock among the species of the Mediterranean lineage of *Squalius* from Eastern Adriatic region are highly congruent with the timing of formation of the Dinaric karst in its current form and tMRCA for all species ranged from Upper Pliocene – Lower Pleistocene (Fig. 1). The oldest estimated divergence among this set of *Squalius* species

is that separating *S. illyricus* from *S. svallize* + *S. zrmanjiae* approximately 2.5 Ma (HPD 95%: 1.2-4.3 Ma). The divergence between *S. microlepis* and *S. tenellus* was very young, close to have occurred one million year ago (HPD 95% 0.4-1.9 Ma), indeed nuclear relationship isolating these two species were not well supported (pp=0.5). This last clade diverged from the Italian *S. lucumonis* at 3.3 Ma (HPD 95% 1.8-5.4 Ma). Land connections between southern Italy and the Balkan-Dalmatian region in ancient times have been suggested (Pasa, 1953). This landbridge from Pliocene up to Pleistocene was several times interrupted and then re-established and would explain relation of Dalmatian *Squalius* species with the Italian *S. lucumonis* as well as the divergence between these species when the landbridge was broke up.

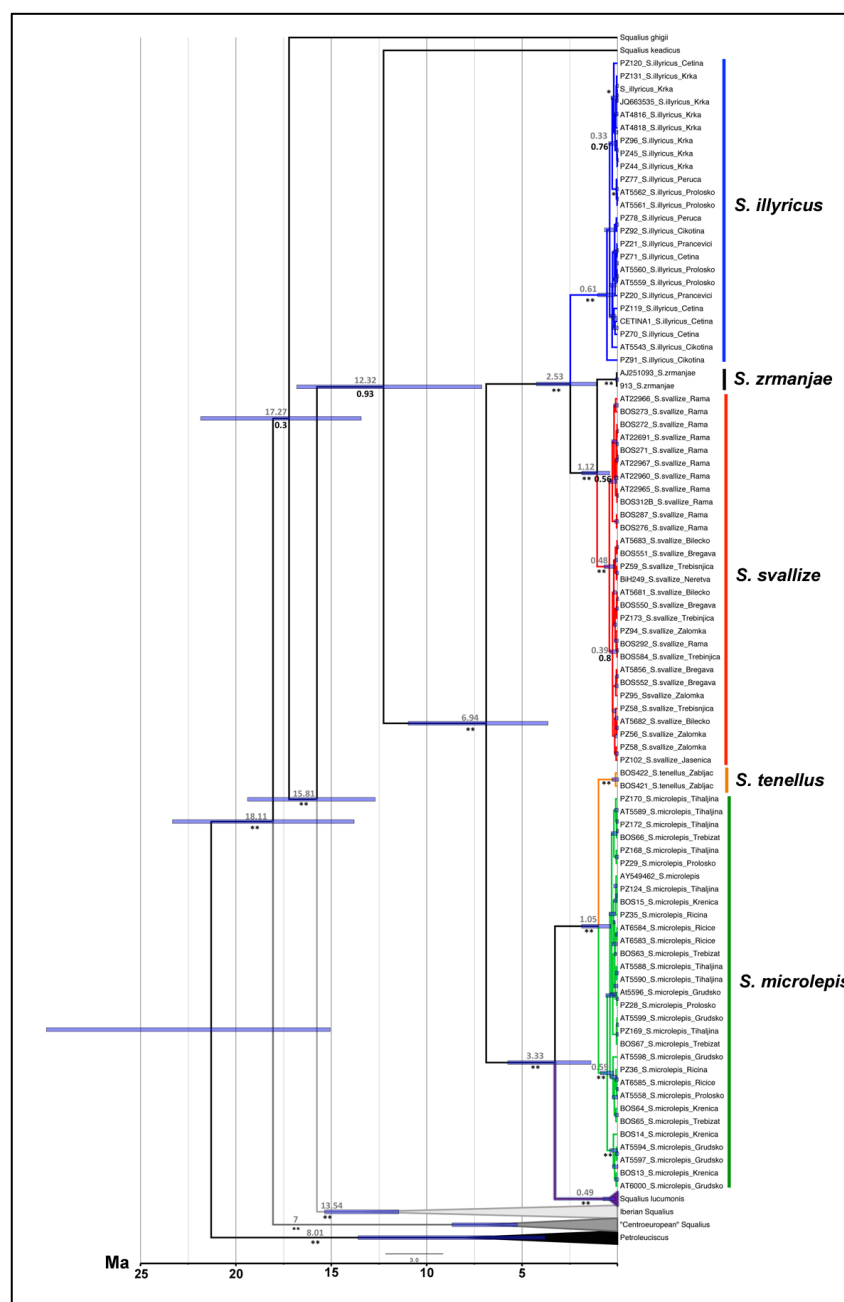


Fig. 1. Bayesian topology and divergence time estimation for the *MT-CYB* gene.

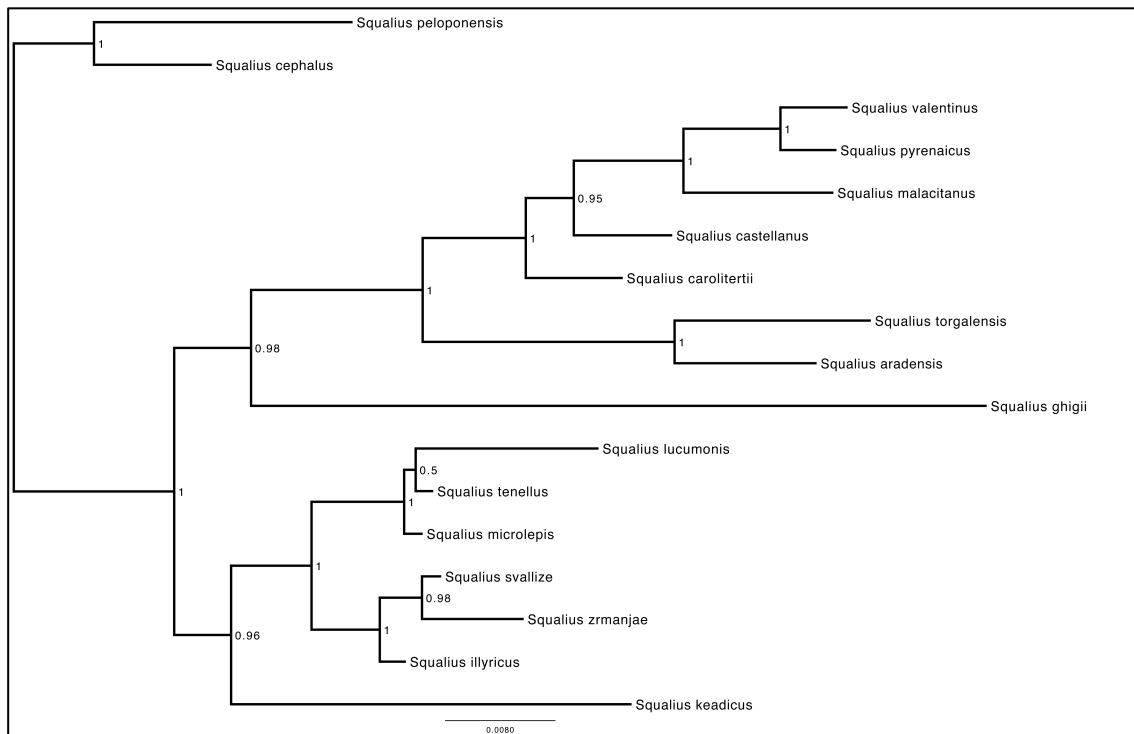


Fig. 2. Bayesian topology for concatenated nuclear genes (RAG1, S7, EFA1α). On nodes posterior probability values.

The young divergence time estimations for the separation between these species from the Eastern Adriatic slope as well as the strong statistical support for mitochondrial and nuclear relationships found in Bayesian topologies suggest that speciation process have been driven by the formation of this geomorphological unit. Similar divergences according to the formation of the Dinaric karst are found in some groups of invertebrates and it has been considered to lead speciation processes for these taxonomic groups (Sket, 1999; Mihevc et al., 2010; Bilandžija et al, 2013; Klobucar et al., 2013; Previsic et al., 2014).

On the other hand, at intraspecific level, our results do not support a strong phylogeographical structure for none of the three *Squalius* species analyzed (*S. illyricus*, *S. svallize* and *S. microlepis*), possibly as a consequence of the recent speciation processes suggested before or possibly as consequence of the influence of the geomorphology of Dinaric karst, with a complex hydrogeographical structure (Bonnaci, 2013).

The species *Squalius svallize* inhabits in Neretva, Trebisnjica and Ljuta basins (Kottelat and Freyhof, 2007), although its distribution and taxonomic status is still doubt (Mrkovcic et al., 2006; Caleta et al., 2015). We have included here populations from these basins plus populations from an additional locality, the endhorreic Zalomka Basin. In Ljuta River we only capture individuals recognized as *Squalius squalus*, a species belonging to the Centroeuropean lineage of this genus. The *S. svallize* populations did not exhibit a marked intraspecific structure, as

showed phylogenetic tree topology and mitochondrial haplotype network. Nevertheless, one tributary of the upper reaches of the Neretva Basin, which currently forms a reservoir in the right margin of the river (Rama Lake), formed a distinct haplogroup in the mitochondrial network (Fig. 3). The lack of phylogenetic structure among Neretva, Trebisnjica and Zalomka, which showed one common haplotype, may be explained by subterranean connections in this karstic region, as some of these rivers, such as Trebisnjica and Neretva, belongs to the same hydrological basin drainage. In fact, it is very common that some species spend part of the time underground (Popovic, 1984; Padalancic et al., 2012a,b) and hence, the movement of individuals among these basins may be feasible, as occur in other Dalmatian freshwater fish species and macroinvertebrates (Trontelj et al., 2007; Padalancic et al., 2012a,b).

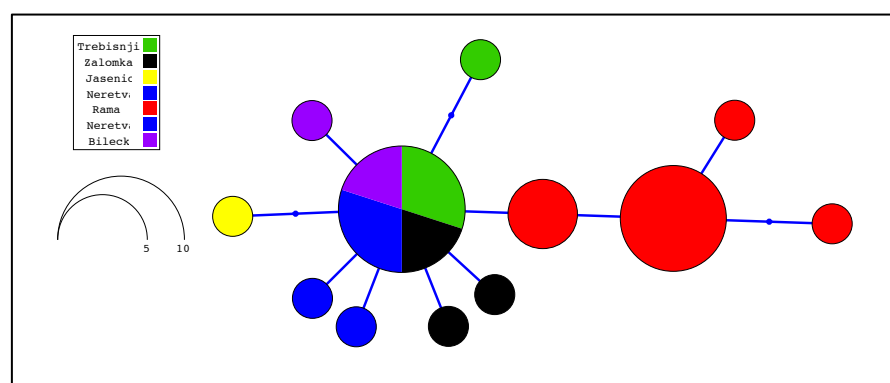


Fig. 3. Mitochondrial haplotype network for the species *S. svallize*

The differentiation of the Rama Lake population is evident, not only by the presence of not shared *S. svallize* haplotypes but also by the presence of some haplotypes of individuals belonging to the Danubian “*Squalius cephalus* complex” (own data not shown) not found in the other three basins. The presence of these Danubian haplotypes probably responds to the fact that the Neretva Basin has two different geological histories. Nevertheless, connections of Danubian tributaries with the upper part of the Neretva basin cannot be discarded, as there is evidence of multiples connections between the Danube Basin and Adriatic river systems throughout the Pleistocene (Habic, 1989; Verovnik et al., 2004).

The species *Squalius illyricus* also shows a lack of phylogenetic and phylogeographic structure in some populations, such is the case of the Cetina Basin (including its tributaries Cikotina and Prancevici) and the Peruca Lake (Fig. 4). However, the Krka basin (Krka river and Roski slap populations) formed an independent group not sharing haplotypes with other populations. Specimens of *Squalius squalus* from Krka basin do not either shared haplotypes with other populations (own data not published). These two concordant phylogeographical patterns in Krka (*S. illyricus* and *S. squalus*) suggest the isolation of this basin and the existence of a disruption in the gene flow among Krka Basin and the remaining Dalmatian *Squalius* populations at some point in the past. The Krka River has been included in the same

biogeographical province than the Po River and other drainages of the North Adriatic Basin (Bianco, 1987, 1995) and an interchange of freshwater fishes from North Italy and Dalmatian, region during the expansion of the Po Basin during the LGM (Waelbrock *et al.*, 2002), as demonstrated phylogenetic relationships of other species belonging to the genus *Scardinius* (Valic *et al.*, 2013).

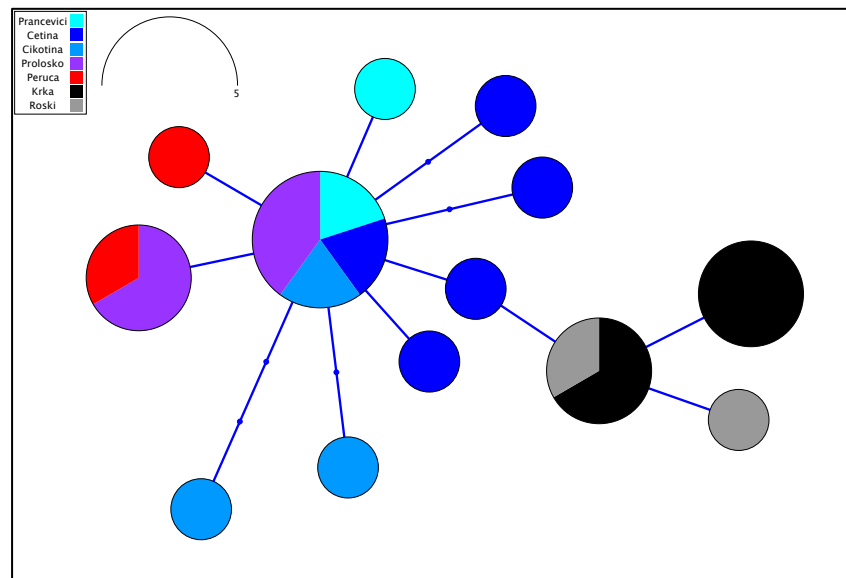


Fig. 4. Mitochondrial haplotype network for the species *S. illyricus*

It is significant the presence of *S. illyricus* haplotypes in the Prolosko lake, which constitutes part of the distribution range of *S. microlepis*. Subterranean connections between Prolosko and other basins may be the main explanation for the presence of *S. illyricus* haplotypes, indeed, there are not private haplotypes in the Prolosko lake. Nonetheless, other reticulated processes such as hybridizations must be discarded.

In turn, the species *Squalius microlepis* occupies a distribution range that covers the area from Prolosko Lake to the rivers Matica, Ricice, Tihaljina y Trebizat (a tributary of the Neretva basin) (Kottelat and Freyhof, 2007; Caleta *et al.*, 2015). This area forms a karstic wetland in which the presence of springs is very common. *S. microlepis* is probably associated to those springs, which is evident from the different haplotype network configuration of this species in relation to *S. illyricus* and *S. svallize* (a larger number of haplotypes but in low frequency) taking into account similar sample sizes. Nevertheless, the lack of phylogenetic and phylogeographic structure and the presence of one haplotype common to all populations as well also point out the great interconnectivity among all these basins inhabited by *S. microlepis*. The presence of these low frequent and private haplotypes may be also associated to these karstic springs, as has been reported for species of the genus *Squalius* from other Mediterranean areas with the presence of coastal karstic springs (Perea *et al.*, 2015).

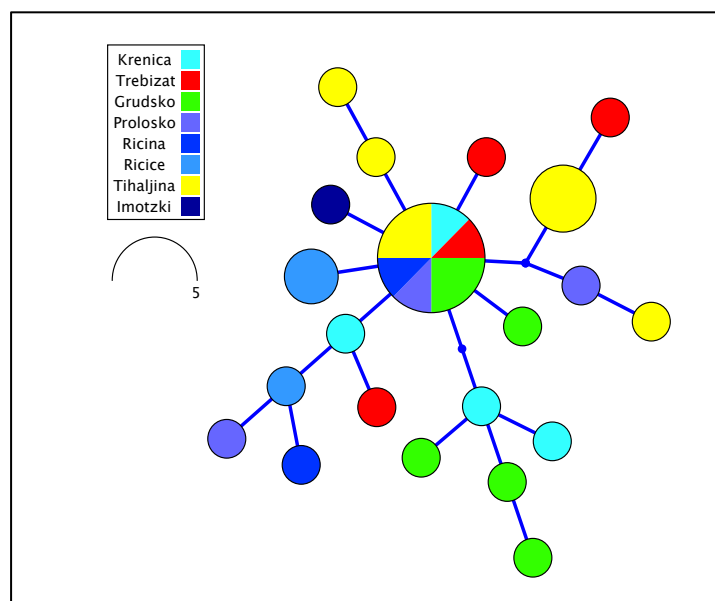


Fig. 5. Mitochondrial haplotype network for the species *S. microlepis*

CONCLUSIONS

Preliminary results obtained in this incipient study suggested a key role of the Dinaric karst upon the speciation processes observed in the *Squalius* species belonging to the Mediterranean evolutionary lineage that inhabit the Eastern. Divergence time estimations for species divergences from this region took place around the Upper Pliocene – Lower Pleistocene, as indicated well-supported mitochondrial and nuclear topologies. Haplotype networks for the three analysed species differ in their structure although in all cases sharing of haplotypes among basins were frequent, especially in the species *S. microlepis*. However, *S. svallize* and *S. illyricus* revealed well-defined geographical haplogroups. These results suggest complex relationships between Adriatic hidrological basins, which can be interconnected underground, and therefore give rise to unique evolutionary processes that are not found in other regions of the world, taking into account the small percentage of this kind of habitats at global level. For this reason, these preliminary results open interesting questions for further specific studies.

REFERENCES

- Assad FA, Jordan H. 1994. Karst terranes and environmental aspects. *Environmental Geology*, 23: 228-237.
- Banarescu PM. 2004. Distribution pattern of the aquatic fauna of the Balkan Peninsula. In: Griffiths HI, Kryštufek B, Reed JM (eds). *Balkan biodiversity-pattern and process in the European hotspot. Part 2*. Dordrecht, The Netherlands: Kluwer Academic Publishers, 203-217.
- Bonacci O, Želković, Galić A. 2013. Karst rivers' particularity: an example from Dinaric karst (Croatia/Bosnia and Herzegovina). *Environmental Earth Sciences*, 70(2): 963-974.
- Bosák P. 2008. Karst processes and time. *Geologos*, 14: 19-36.
- Buj I, Vukić J, Šanda R, Perea S, Čaleta M, Marčić Z, Bogut I, Povz M, Mrakovčić M. 2010. Morphological

- comparisons of bleaks (*Alburnus*, Cyprinidae) from the Adriatic Basin with the description of a new species. *Folia Zoologica*, 59(2): 129-141.
- Buj I, Šanda R, Marčić Z, Čaleta M, Mrakovčić M. 2014. Combining morphology and genetics in resolving taxonomy – a systematic revision of spined loaches (genus *Cobitis*; Cypriniformes, Actinopterygii). *PLoS One*, 9: e99833.
- Čaleta M, Buj I, Mrakovčić M, Mustafić P, Zanella D, Marčić Z, Duplić A, Mihinjač T & Katavić I. 2015. Endemic Fishes of Croatia. Croatian Environment Agency. Zagreb. 116 pp.
- Cohen AS. 2003. Paleolimnology: The history and evolution of lake systems. Oxford University Press. New York. 500 pp.
- Coissac R, Riaz T, Puillandre N. 2012. Bioinformatic challenges for DNA metabarcoding of plants and animals. *Molecular Ecology*, 21: 1834-1847.
- Freyhof J, Brooks, 2011. European red list of freshwater fishes.
- Habic. 1989. Karstic bifurcations of Pivka river on the Adriatic-Danubian watershed. *Acta Carsologica*, 18: 233-264.
- Higgins D, Thompson J, Gibson T, Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acid Research*, 22: 4673-4680.
- Klobučar GIV, Podnar M, Jelić M, Franjević D, Faller M, Štambuk A, Gottstein S, Simić V, Maguire I. Role of the Dinaric karst (western Balkans) in shaping the phylogeographic structure of the threatened crayfish *Austropotamobius torrentium*. *Freshwater Biology*. 2013; 58:1089–1105.
- Kottelat M, Freyhof J. 2008. *Squalius tenellus*. The IUCN Red List of Threatened Species. 2008: e.T135695A4184809.
- Lanfear R, Calcott B, Ho SYW, Guindon S. 2012. PartitionFinder: Combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution*, 29(6): 1695-1701.
- Leiner S, Popović J. 1984. Genus *Leuciscus* (Cyprinidae, Pisces) in waters of the Adriatic river basin regarding *L. svallize* find (Heckel & Kner, 1858) and *L. souffia* Risso, 1826 in Cetina. *Ichthyologica*, 16: 111-120. (In Croatian with English summary).
- Mihevc A. 2001. The Speleogenesis of Divača Karst. ZRC. Press. Ljubljana. (In Slovene).
- Palandačić A, Bonnaci O, Snoj A. 2012. Molecular data as a possible tool for tracing groundwater flow in karst environment: example of *Delminichthys adspersus* in Dinaric karst system. *Ecohydrology*, 5(6): 791-797.
- Palandačić A, Matschiner M, Zupančič P, Snoj A. 2012. Fish migrate underground: the example of *Delminichthys adspersus* (Cyprinidae). *Molecular Ecology*, 21(7): 1658-1671.
- Pasa A. 1953. Appunti geologici per la paleogeografia delle Puglie. *Memorie Biogeogr. Adriat.*, 2: 175-286.
- Previšić A, Schnitzler J, Kucinic M, Graf W, Ibrahimi H, Kerovec M, Pauls SU. 2014. Microscale vicariance and diversification of Western Balkan Caddisflies linked to karstification. *Freshwater Science*, 33(1): 250-262.
- Sanda R, Bogut I, Doadrio I, Perdices A, Perea S, Sedivá A & Vukic J. 2008. Distribution and taxonomic relationships of spined loaches (Cobitidae, *Cobitis*) in the River Neretva basin, Bosnia and Herzegovina. *Folia Zoologica*, 57(1-2): 20-25.
- Šlechtová V, Bohlen J, Freyhof J, Persat H, Delmastro GB. 2004. The Alps as barrier to dispersal in cold-adapted freshwater fishes? Phylogeographic history and taxonomic status of the bullhead in the Adriatic freshwater drainage. *Molecular Phylogenetics and Evolution*, 33: 225-239.
- Sket B. 1999. The nature of biodiversity in hypogean waters and now it is endangered. *Biodiversity and Conservation*, 8: 1319-1338.
- Trontelj P, Machino Y, Sket B. 2005. Phylogenetic and phylogeographic relationships in the crayfish genus *Austropotamobius* inferred from mitochondrial COI gene sequences. *Molecular Phylogenetic and Evolution*, 34: 212-226.
- Valić D, Vardić Smrzlić I, Kapetanović D, Teskeredžić, Pleše B, Teskeredžić E. 2013. Identification, phylogenetic relationships and a new maximum size of two rudd populations (*Scardinius*, Cyprinidae) from the Adriatic Sea drainage, Croatia. *Biologia*, 68: 539-545.

Verovnik R, Sket B, Trontelj P. 2005. The colonization of Europe by the freshwater crustacean *Asellus aquaticus* (Crustacea: Isopoda) proceeded from ancient refugia and was directed by habitat connectivity. *Molecular Ecology*, 14: 4355-4369.

Phylogeography, historical demography and habitat suitability modeling of freshwater fishes inhabiting seasonally fluctuating Mediterranean systems: a case study using the Iberian cyprinid *Squalius valentinus*

Perea S & Doadrio I. 2015. Phylogeography, historical demography and habitat suitability modeling of freshwater fishes inhabiting seasonally fluctuating Mediterranean river systems: a case study using the Iberian cyprinid *Squalius valentinus*. *Molecular Ecology*, 24: 3706-3722.

DOI: <http://dx.doi.org/10.1111/mec.13274>

Perea S & Doadrio I. 2015. Phylogeography, historical demography and habitat suitability modeling of freshwater fishes inhabiting seasonally fluctuating Mediterranean river systems: a case study using the Iberian cyprinid *Squalius valentinus*. *Molecular Ecology*, 24: 4808. (Corrigendum)

DOI: <http://dx.doi.org/10.1111/mec.13352>

RESUMEN

Los peces de agua dulce mediterráneos han evolucionado bajo las restricciones impuestas por los patrones estacionales climáticos e hidrológicos que definen el clima Mediterráneo. Estas condiciones mediterráneas han influenciado la estructura genética y demográfica de las comunidades acuáticas desde su origen en el Plioceno Medio. Las especies de agua dulce de regiones con patrones climáticos similares al clima Mediterráneo probablemente constituirán poblaciones bien estructuradas en mayor o menor medida dependiendo del tamaño de las cuencas, como consecuencia de la fragmentación resultante de los ciclos de sequía y lluvias, muchas veces en forma de riadas. En este trabajo se ha desarrollado un marco integrativo para estudiar los patrones espaciales de diversidad genética, tendencias demográficas, modelaje de idoneidad de hábitat y genética del paisaje con el fin de evaluar la respuesta evolutiva de los peces de agua dulce mediterráneos a las fluctuaciones climáticas estacionales. Para testar esta respuesta evolutiva se ha usado la especie *Squalius valentinus* como modelo de estudio, una especie de ciprínido endémica del levante español, donde las fluctuaciones estacionales pueden ser extremas. Sin embargo, nuestros hallazgos pueden ser extrapolados a otras especies mediterráneas. Nuestros resultados ponen de manifiesto el significativo papel del clima Mediterráneo, junto con las glaciaciones pleistocénicas, en la diversificación de *S. valentinus*. Nosotros encontramos una diversidad genética nuclear más alta en las cuencas hidrológicas más grandes, pero la diversidad genética mitocondrial estuvo más relacionada con la idoneidad de hábitat que con el tamaño de la cuenca. También encontramos una fuerte correlación entre la estructura genética y los factores climáticos asociados a la estacionalidad mediterránea. Los análisis demográficos y de migración sugieren una expansión poblacional durante los periodos glaciares que también contribuyó a la estructura genética actual de las poblaciones de *S. valentinus*. Los modelos inferidos apoyan la contribución significativa de la precipitación y la temperatura a la de idoneidad de hábitat de *S. valentinus* y permiten reconocer áreas de estabilidad de hábitat. Nosotros destacamos la importancia de condiciones de hábitat estables favorecidas por los típicos surgimientos de agua kársticos de las zonas costeras litorales mediterráneas para la conservación de las especies de agua dulce que habitan ríos estacionalmente fluctuantes.

Palabras clave: Demografía histórica, Filogeografía, Modelos de distribución, Peces de agua dulce, Clima Mediterráneo, Ríos estacionalmente fluctuantes.

ABSTRACT

The Mediterranean freshwater fish fauna has evolved under constraints imposed by the seasonal weather/hydrological patterns that define the Mediterranean climate. These conditions have influenced the genetic and demographic structure of aquatic communities since their origins in the Mid-Pliocene. Freshwater species in Mediterranean-type climates will likely constitute genetically well-differentiated populations, to varying extents depending on basin size, as a consequence of fragmentation resulting from droughts/flood cycles. We developed an integrative framework to study the spatial patterns in genetic diversity, demographic trends, habitat suitability modelling and landscape genetics, to evaluate the evolutionary response of Mediterranean-type freshwater fish to seasonal fluctuations in weather. To test this evolutionary response, the model species used was *S. valentinus*, an endemic cyprinid of the Spanish Levantine area, where seasonal weather fluctuations are extreme, although our findings may be extrapolated to other Mediterranean-type species. Our results underscore the significant role of the Mediterranean climate, along with Pleistocene glaciations, in diversification of *S. valentinus*. We found higher nuclear diversity in larger drainage basins, but higher mitochondrial diversity correlated to habitat suitability rather than basin size. We also found strong correlation between genetic structure and climatic factors associated with Mediterranean seasonality. Demographic and migration analyses suggested population expansion during glacial periods that also contributed to the current genetic structure of *S. valentinus* populations. The inferred models support the significant contribution of precipitation and temperature to *S. valentinus* habitat suitability and allow recognizing areas of habitat stability. We highlight the importance of stable habitat conditions, forested by typical karstic springs found on the Mediterranean littoral coasts, for preservation of freshwater species inhabiting seasonally fluctuating river systems.

Keywords: Distribution modelling, Freshwater fishes, Historical demography, Mediterranean climate, Phylogeography, Seasonally fluctuating rivers.

INTRODUCTION

In the Mediterranean basin, the southern European peninsulas have been recognized as the primary refugia for many species during the Late Cenozoic, most showing high genetic diversity resulting from isolation of genotypes and gain or loss of alleles (Hewitt, 2004). The importance of the peninsulas in this respect has been widely debated for flora and terrestrial fauna (Hewitt, 2004; Provan and Bennet, 2008; Sommer *et al.*, 2009; Fuentes-Utrilla *et al.*, 2014), while there has been little interest in aquatic fauna (e.g. Gómez and Lunt, 2007; Gante *et al.*, 2009). Nevertheless, aquatic organisms are good models for study of historical events that shaped current species genetic diversity and structure, especially those organisms with low dispersal opportunities such as freshwater fishes.

Since the end of the Pliocene and throughout the Pleistocene, the most notable historical factors affecting genetic diversity and structure of aquatic organisms in the Mediterranean basin include the formation of the hydrological network, periodic shifts in shorelines due to marine transgressions and regressions, and the development of the Mediterranean climate. These factors likely exerted a drastic influence on the genetic diversification patterns of aquatic organisms and could contributed substantially to explaining the high genetic diversity found in southern European refugia (Hewitt, 2004).

The onset of the Mediterranean climate occurred in the Mid-Pliocene period, when summers started become arid (Fauquette *et al.*, 1999; Jiménez-Moreno *et al.*, 2010). During the Late Pliocene, a change of climatic conditions occurred, culminating in the highly variable climatology of the Pleistocene glacial/interglacial cycles (Haywood *et al.*, 2009). In general, glacial/interglacial cycles have been considered responsible for extensive modification in fluvial dynamics in southern European Mediterranean peninsulas (Hughes *et al.*, 2006). This fluvial modification, together with seasonal fluctuations, could have played an important role in the genetic and demographic structure and variability of strict freshwater fish fauna since the Plio-Pleistocene period.

The Mediterranean climate is characterized by periods of drought alternating with rainfall that may occasionally be intense and even catastrophic, especially affecting smaller and shorter river systems (Benito *et al.*, 2008; Boix *et al.*, 2010). Drought periods promote fragmentation events as a consequence of the reduction in hydrological connectivity among rivers, and among stretches of a given river, creating isolated summer pools. In the long term, this determines freshwater population structure and demography (Blondel and Aronson, 1999; Humphries and Baldwin, 2003), influencing population density and size and inter- and intraspecific relationships

as well as genetic structure and diversity (Davey and Kelley, 2007; Dekar and Macgoulick, 2007). In contrast, flooding restores connectivity, facilitating gene flow among populations (Hughes and Hyllier, 2006). The size and geomorphology of Mediterranean-type basins in which freshwater fish fauna occur are determinants of the magnitude of the effects of seasonal and interannual fluctuations on the genetic and demographic history of their populations. In Mediterranean-type systems, higher genetic structure and lower effective population size are expected in the smaller and shallower river systems, as consequence of more pronounced fragmentation events (Lowe and Allendorf, 2010). Extreme conditions may lead to loss of individuals and, consequently, of genetic diversity (Luikart *et al.*, 2003; Bond *et al.*, 2008). On the contrary, higher effective population sizes are expected in larger and deeper systems and, thus, higher levels of genetic diversity (Frankham, 1996).

Taking into account these Plio-Pleistocene climatic scenarios, we present a study case to evaluate the impact of fluctuating river systems and interaction of glacial cycles with the Mediterranean climate on the evolutionary history of strict freshwater fish fauna from Mediterranean rivers. As a model species, we selected an endemic Iberian fish, *Squalius valentinus*. The distribution range of *S. valentinus*, in the Spanish Levantine region, is coincident with an area in which seasonal and interannual weather fluctuations may occasionally be extreme. The region has a semi-arid climate (Lázaro *et al.*, 2001) and exhibits two distinct river sizes that are differentially affected by seasonal fluctuations: large rivers (156-498 km in length) in the northern area and small basins (10-81 km in length) in the southernmost area, where *S. valentinus* is commonly the only fish species present (Doadrio *et al.*, 2011). The northern basins are large and branched and exhibit two well-differentiated regions, an upper-middle area with a steep profile and a pluvio-nival precipitation regime and a downstream alluvial plain, in which disconnected ponds are common in summer. The southern basins are constituted by small rivers having deep sources that are less impacted by weather fluctuations, and small intermittent rivers. Summer drought periods alternating with sudden discharges of large volumes of water during heavy rainfall may have extreme effects in these smaller basins, which occasionally even remain dry for several years, and hence, their freshwater populations will suffer strong impact.

The cyprinid *S. valentinus* belongs to one of the most diversified genera of freshwater fishes in Europe (Kottelat and Freyhof, 2007), demonstrating its evolutionary success in colonizing varied habitats, among them Mediterranean-type freshwater systems. The Middle Pliocene origin of *S. valentinus* (Doadrio and Carmona, 2006), coincident with the appearance of summer droughts (Fauquette *et al.*, 1999; Jiménez-Moreno *et al.*, 2010), along with its current distribution range covering most of the Spanish Levantine region, offers the opportunity to assess the evolutionary response of a primary fish species to seasonally fluctuating weather conditions.

This evolutionary response will be reflected in the current genetics and historical demography of their populations. We hypothesised that genetic structure and diversity of freshwater species inhabiting Mediterranean-type systems were strongly influenced by fluctuating climatic conditions of the Plio-Pleistocene period and that the effects of these seasonal fluctuations on genetic diversity and demography are stronger in populations of smaller and shallower river systems than in those inhabiting larger and deeper systems.

To test these hypotheses, we selected *S. valentinus* for a case study to achieve the following aims: (i) to reveal the mitochondrial and nuclear genetic structure of populations and to determine its spatial configuration, (ii) to investigate the demographic history of its populations and (iii) to infer, through species distribution models and landscape genetic approaches, the main bioclimatic factors that explain *S. valentinus* evolutionary history.

MATERIAL AND METHODS

1. Sampling, DNA extraction, amplification and sequencing

Specimens of *Squalius valentinus* were collected by electrofishing, with permission from the relevant authorities, across most of its distribution range (Fig. 1; Table S1 in the Supporting Information of this chapter). Thirty specimens per population were examined for both mitochondrial and nuclear genes when possible. Total genomic DNA was extracted from fin-clip tissue preserved in ethanol 95% using the commercial kit Biosprint15 for tissue and blood (Qiagen). For each specimen, a 1098-bp fragment of the mitochondrial cytochrome *b* (*MT-CYB*) gene and a 1086-bp fragment of the exon 3 of the single copy protein-encoding nuclear recombination-activating gene 1 (*RAG1*) were amplified. *MT-CYB* has proven useful in phylogenetic and phylogeographic studies of freshwater fishes (Ornela-García *et al.*, 2008; Doadrio *et al.*, 2009) and in Iberian cyprinids in particular (Doadrio and Carmona, 2003, 2006; Sanjur *et al.*, 2003). The use of *RAG1* has been validated in phylogenetic and phylogeographic studies of freshwater fish (Domínguez-Domínguez *et al.*, 2011), among them the genus *Squalius* (Perea *et al.*, 2010). Primers and protocols used for PCR products on 1% agarose gels, they were purified by ExoSAP-IT (USB) and directly sequenced by Macrogen using a 3730XL DNA sequencer.

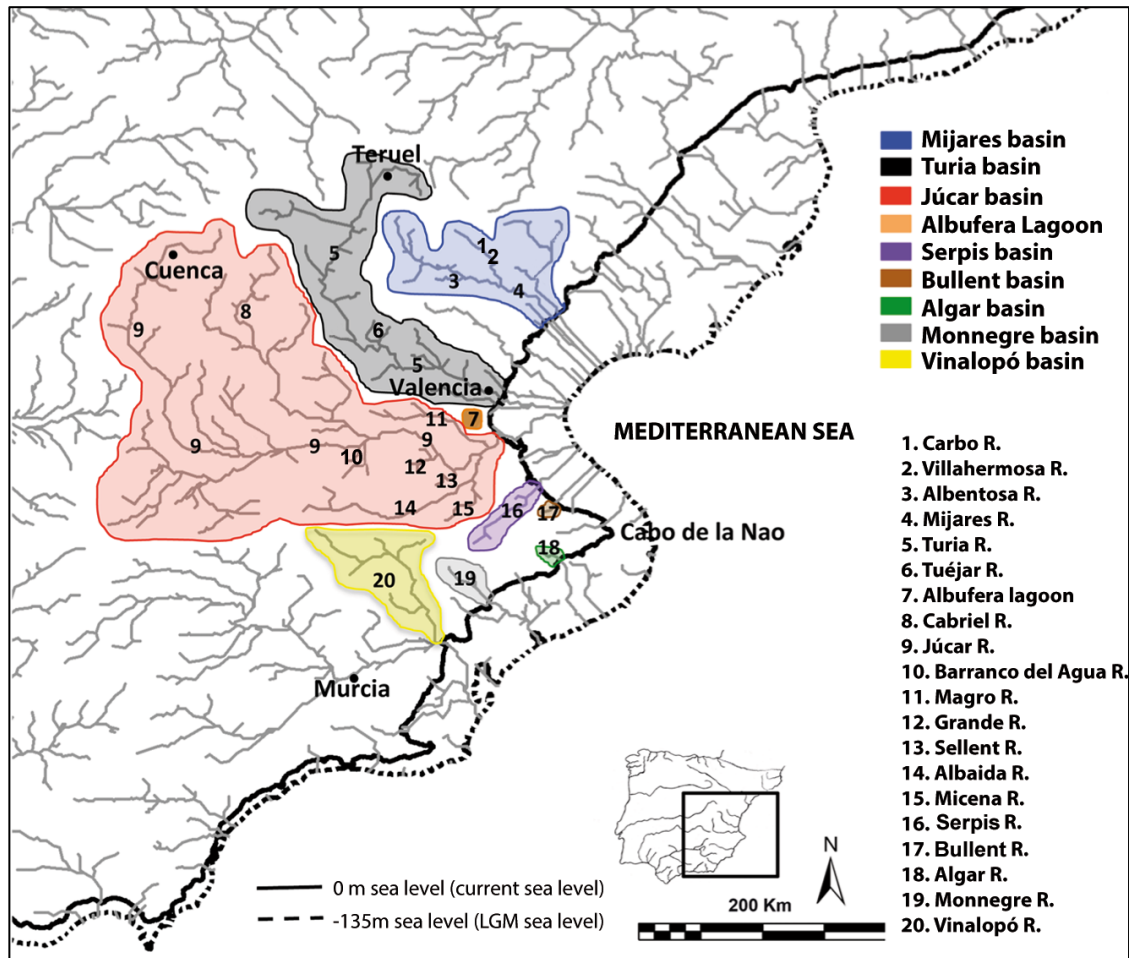


Fig. 1. Geographical distribution of *Squalius valentinus* populations, sampling localities and historical bathymetric predictions during the Last Glacial Maximum (LGM) sea-level retraction in the Spanish Levante area.

2. Phylogeography, genetic diversity and genetic structure

To test the effects of seasonal weather fluctuations on genetic diversity and structure, and its consequences in association with size and geomorphology of rivers, we conducted a phylogeographic analysis.

All sequences were initially aligned using the default pairwise and multiple alignment parameters in Clustal W (Thompson *et al.*, 1994) implemented in MEGA v.5.1 (Tamura *et al.* 2011). Chromatograms and alignments were later revised. For nuclear data, alleles were unphased according to PHASE algorithm implemented in DNAsp v.5.0 (Librado and Rozas, 2009). The pairwise homoplasmy index (Bruent *et al.*, 2006) to detect genetic recombination in the nuclear RAG1 gene was calculated by the software SPLITSTREE v.4.10 (Hudson & Bryant,

2006). Significant evidence of recombination was not found ($p=0.33$). Haplotype and allele relationships were obtained using HaploViewer (<http://www.cibiv.at/~%20greg/haploviewer>).

To characterize genetic diversity for each population, haplotype and nucleotide diversity were estimated using DNAsp v.5.0 (Librado and Rozas, 2009). Spatial distribution of haplotype and nucleotide diversity was summarized using the Spatial Analyst extension in ARCMAP v.9.3 (ESRI, Redlands, CA, USA) and the ordinary Kriging method. Genetic differentiation among populations was assessed in ARLEQUIN v.3.5 (Excoffier and Lischer, 2010) using the fixation index F_{ST} (Weir and Cockerham, 1984). The significance of F_{ST} pairwise comparisons was given by a p -value calculated using 10^4 random permutation tests. P -values were further adjusted by Bonferroni's correction (Rice, 1989). Evidence of genetic hierarchical population structure was assessed with analysis of molecular variance (AMOVA) for both genes, implemented in ARLEQUIN v.3.5 (Excoffier and Lischer, 2010). We first examined the overall differences of the complete data set (one gene pool) and then determined the partitioning of genetic variation, treating each basin separately (nine gene pools). Global genetic structure of *S. valentinus* was investigated for both genes using spatial analysis of molecular variance (SAMOVA) (Dupaunloop *et al.* 2002) run from $K=1$ to $K=20$ with 500 simulated annealing processes. Each run was conducted in triplicate.

3. Population size changes and migration rate

To evaluate the influence of seasonal fluctuations on historical demography of *S. valentinus*, deviations from a model of mutation-drift equilibrium were tested for the *MT-CYB* gene using Fu's F_s (Fu, 1997) and R_2 (Ramos-Onsins and Rozas, 2002) with 1000 pseudo-replications, as well as with Tajima's D (Tajima, 1989) neutrality tests. Tests were conducted using DNAsp v.5.0 (Librado and Rozas, 2009). To visualize historical demographic trends, we conducted mismatch analyses in DNAsp v.5.0 (Librado and Rozas, 2009). The fit of the data to the sudden demographic expansion model was assessed by the probability of obtaining smaller raggedness values (r) (Rogers and Harpending, 1992) than those observed under coalescent algorithm simulations over 1000 pseudo-replications and with no recombination. Initial values were set at $\theta_0=0$ and $\theta_1=99,999$. When neutrality test values were significant, a rough estimate was made of the time (t) corresponding with population expansion events found by mismatch analyses, using the formula $t = \pi/2(k\mu)$ (Slatkin and Hudson, 1991). Calculation of mutation rate (μ) used a generation time of 2 years for *S. valentinus* based on that observed in other Iberian cyprinids of similar size occupying similar habitats (Magalhães *et al.* 2003). The evolutionary rate used was 0.4% per million years per lineage (Perea *et al.*, 2010).

An extended Bayesian skyline plot (EBSP) to reveal demographic changes over time under neutral evolution (Drummond *et al.*, 2005) was implemented in BEAST v.1.7 (Drummond and Rambaut, 2007). EBSP allows inclusion of mitochondrial and nuclear loci in the same analysis. Nuclear data were unphased to allow inclusion in the analysis. The EBSP was performed for the entire data set and for each basin independently, applying the same evolutionary rate as mismatch analyses for *MT-CYB*. The evolutionary rate of the *RAG1* gene was estimated as a function of the mitochondrial evolutionary rate. An uncorrelated molecular clock model was set as prior, and 100^6 generations were run in each analysis. Convergence was assessed with TRACER v.1.5 (Rambaut and Drummond, 2007).

To assess connectivity among populations, historical migration rates, along with their magnitude and direction, based on *MT-CYB* were computed using MCMC simulations in MIGRATE-N (Beerli, 2009). MCMC searches were conducted with one long chain and 12 short parallel chains with start value temperatures (1.0, 1.5, 3.0, 100,000.0) and adaptive-bound heating. Final MCMC searches used 2,500,000 recorder steps in 100 increments, discarding the first 10^4 generations as burn-in. Initial parameters were obtained from π_{ST} estimates. We set an initial uniform prior as μ (0.0-0.1) and M (0.0-25,000.0).

4. Species distribution modelling and landscape genetic approaches

To evaluate the influence of historical sea-level changes on *S. valentinus* population structure and demographics, bathymetric data were obtained from a 30 arc-second (c. 1 km) data set from GEBCO 08 (September, 2010, www.gebco.com). GEBCO data were transformed to raster format, and the raster file was modified to produce historical [Last Glacial Maximum (LGM); 21,000 years ago] and current bathymetric contours using ARCINFO hydrological tools (ARCMAP and ARCINFO v.9.3; ESRI).

To evaluate the main contributory bioclimatic variables to habitat suitability for *S. valentinus*, we carried out species distribution modelling at different temporal scales (Elith *et al.*, 2006). With this approach, sites with environments similar to those in which species has been observed are identified as potential occurrence areas. Estimates of the current and past population distribution were constructed with MAXENT v.3.0 (Phillips *et al.*, 2006). Geographical coordinates of the 36 sampling sites (Table S1 in the Supporting Information of this chapter) along with sites accessed in the Global Biodiversity Information Facility database (693 occurrences summarized to 78 points; www.gbif.es) were used as presence data (114 sites in total; Table S2 in the Supporting Information of this chapter). We used default settings, and the model was run with 100 subsample replicates, estimating mean habitat suitability values (S). The

climate model was based on the WORLDCLIM database (Hijmans *et al.*, 2005; <http://www.worldclim.org>). The WORLDCLIM variables represent biologically meaningful summaries of precipitation and temperature in the present (1950-2000), the LGM [Community Climate System Model (CCSM)], and the last interglacial (LIG; 140,000-120,000 years ago) periods. Future distribution range was inferred using the IPCC4 data set from CIAT (<http://www.ccafsclimate.org/data/>) with the delta change method (Reynard *et al.*, 2001), Ab1 scenario (moderate emission; Nakicenovic and Swart, 2000), and the CCSM. Accuracy of all distribution models was assessed using 75% of the presence data to train the model and the remaining 25% for testing. A jackknife test of variable importance was conducted to evaluate the relative strength of each climate variable (Pearson *et al.*, 2007). Variables contributing least to the model or showing multicollinearity (Graham, 2003) were removed (see Appendix S1 in the Supporting Information of this chapter). The area under the receiver operating curve (AUC) was estimated to assess the accuracy of the model. With the aim to assess changes in habitat suitability along time throughout niche overlapping, we selected two thresholds to categorize the potential distribution of *S. valentinus* in “suitable” and “unsuitable” habitats. The chosen thresholds were the *minimum training presence logistic threshold* and the *10th percentile training presence logistic threshold* commonly used for models comparison (e.g. Pearson, 2007; Rödder *et al.*, 2009). Model outputs were manipulated in ARCMAP v.9.3 (ESRI) in order to reclassify them into binary suitable/unsuitable models according to these two thresholds. Niche overlapping was also calculated in ARCMAP.

After obtaining habitat suitability models (based on bioclimatic variables) for all periods, general linear modelling (GLM) was used to assess its effects on the genetic variability (diversity) and genetic variation (structure) of *S. valentinus* populations over time. Also using GLM, we tested the individual interactive effects of past and present-day bioclimatic variables (Hijmans *et al.*, 2005) on the genetic structure of *S. valentinus*. To assess the effect of habitat suitability on genetic diversity, GLM was conducted in STATISTICA v.5.1 (StatSoft Inc.). For genetic structure, multiple regression was carried out using the hierarchical Bayesian method implemented in GESTE program (Foll and Gaggiotti, 2006). We analysed different sets of variables to estimate which kind of environmental factors exerted strong effect on the genetic structure of *S. valentinus* populations: temperature (BIO1-11), precipitation (BIO12-19), seasonality (BIO4, BIO7 y BIO15) or “extreme” climatic conditions (BIO5, BIO6, BIO13 and BIO14). Description of variables is given in Appendix S1 (Supporting Information of this chapter). This approach also allowed dealing with multicollinearity by reducing the number of variables included in each analysis. For temperature and precipitation sets, we conducted a second analysis including only the five factors that best explained the observed genetic structure, which is the five variables with the highest posterior probabilities after summing the probability from all models that included the variable (Gaggiotti *et*

al., 2009). We used the same run parameters for this second analysis. As genetic data input, we identified 63 single nucleotide polymorphisms (SNPs) in the *MT-CYB* gene. For environmental data input, we extracted the values of bioclimatic map layers (bioclimatic factors) and MAXENT modelling output (habitat suitability factor) for each sampling site, using ARCGIS v.9.3 software. The Foll and Gaggiotti method estimates population-specific F_{ST} values for each local population and relates them to environmental factors using GLMs. This method also provides a posterior probability value for each of the models (reversible jump MCMC). We used 10 pilot runs of 5000 iterations, a burn-in of 5×10^6 iterations, a thinning interval of 50 and a sample size of 50,000. All models were evaluated by the 95% highest probability density interval.

RESULTS

1. Phylogeography, genetic structure and genetic diversity

A fragment of 1098 bp for *MT-CYB* and a fragment of 1086 bp for the nuclear *RAG1* gene were sequenced from a total of 478 and 381 (762 alleles), respectively. Forty-nine mitochondrial haplotypes and 28 nuclear alleles were identified (Table S3, in the Supporting Information of this chapter). Mitochondrial and nuclear networks (Fig. 2A,B) revealed star-shaped structure with a single central haplotype and allele shared by all populations (H1=254, A1=633). In nearly all basins, as well as in rivers within basins, unique haplotypes and alleles were observed, sometimes at low frequencies (Fig. 2A,B).

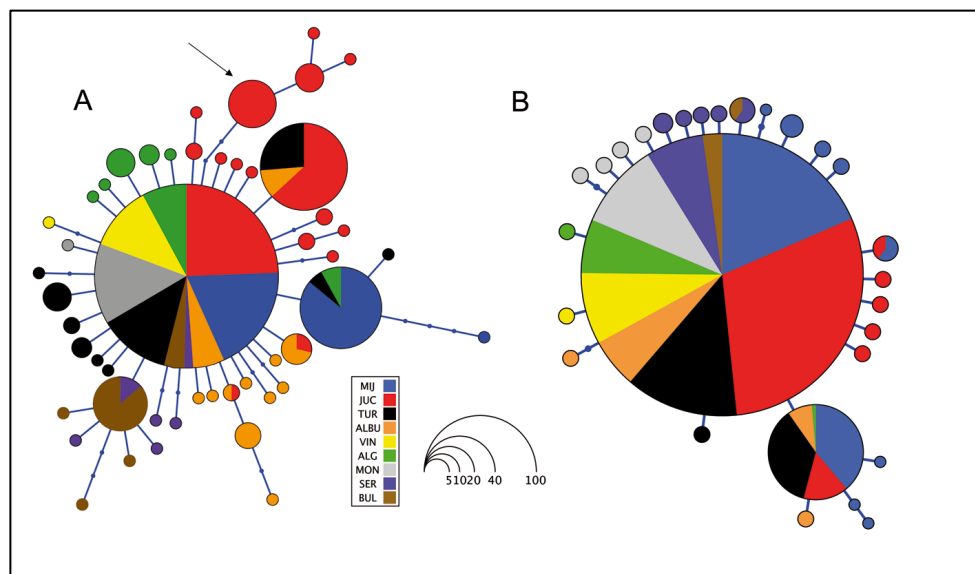


Fig. 2. Haplotype networks of mitochondrial cytochrome b gene (A) and nuclear *RAG1* gene (B) for *Squalius valentinus* populations. The size of each haplotype symbol is proportional to its frequency. Black arrow identifies the haplogroup of downstream Júcar.

Downstream areas of the larger northern basins and the smaller southern basins with deeper river stretches near their source showed the highest mitochondrial haplotype and nucleotide diversity (Table S3 in the Supporting Information of this chapter; Fig. 3). Within the northern basins, remarkably genetic diversity values were higher in downstream areas. In this area, remarkably high genetic diversity values were found in a freshwater coastal lagoon, the Albufera. Patterns of nuclear genetic diversity differed slightly from mitochondrial (Fig. 3), showing a descending gradient in genetic diversity from the northern to the southern distribution range. This result is congruent with the significant correlation found between genetic diversity and basin size in nuclear but not in mitochondrial DNA (Table 1).

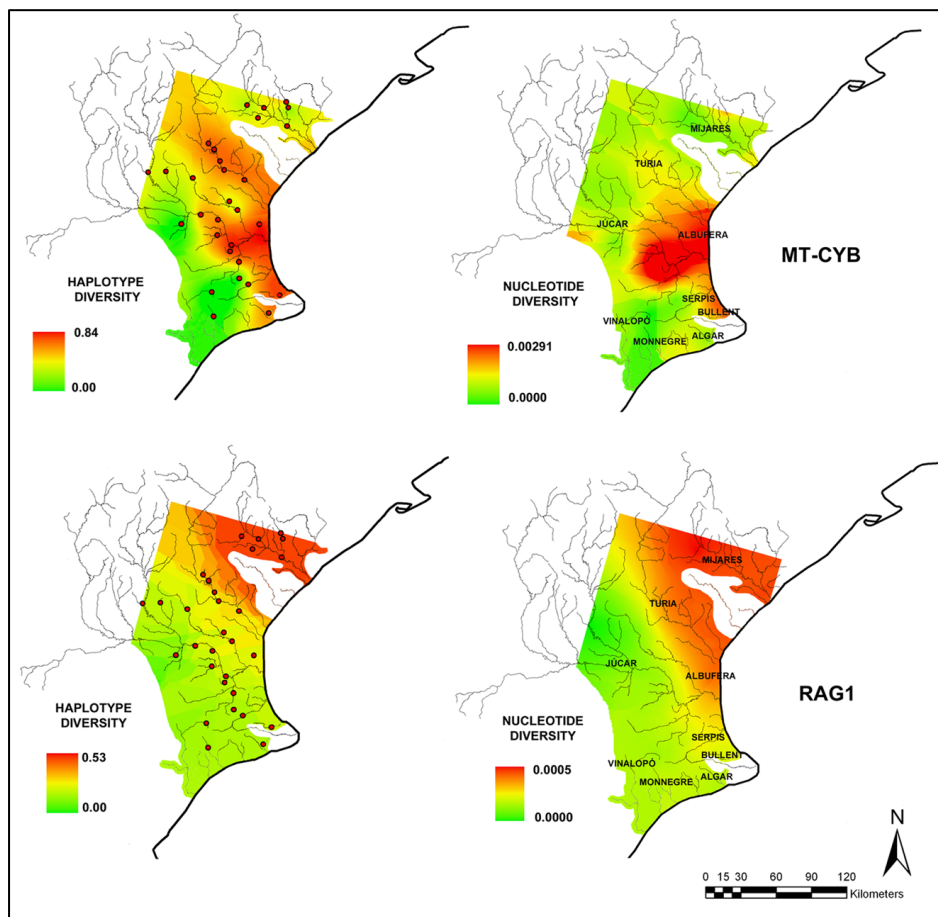


Fig. 3. Estimated distribution of haplotype and nucleotide diversity across the distribution range of *Squalius valentinus*. (A) Mitochondrial cytochrome b. (B) Nuclear RAG1 gene. Red indicates higher values of genetic diversity, while green represents lower values.

Table 1. Spearman correlation (r) between genetic diversity and basin size (left) and habitat suitability (right). BL, Basin length; BS, Basin surface; MtHD, Mitochondrial haplotype diversity; MtPi, Mitochondrial nucleotide diversity; NucHD, nuclear haplotype diversity; NucPi, nuclear nucleotide diversity; SLIG, habitat suitability for LIG; SLGM, habitat suitability for LGM; SPD, habitat suitability for Present-day; S2080, habitat suitability for future; LIG, Last Interglacial; LGM, Last Glacial Maximum. * Significant *p*-level values.

Genetic diversity vs basin size			Genetic diversity vs habitat suitability		
Correlation	r	p	Correlation	r	p
r(BL/MtHD)	0.16	0.68	r(SLIG/MtHD)	0.45	0.04*
r(BL/MtPi)	0.12	0.75	r(SLIG/MtPi)	0.60	0.005*
r(BL/NucHD)	0.73	0.02*	r(SLIG/NucHD)	0.44	0.014*
r(BL/NucPi)	-0.26	0.50	r(SLIG/NucPi)	0.42	0.018*
r(BS/MtHD)	0.20	0.61	r(SLGM/MtHD)	0.47	0.03*
r(BS/MtPi)	0.19	0.62	r(SLGM/MtPi)	0.39	0.03*
r(BS/NucHD)	0.70	0.03*	r(SLGM/NucHD)	0.19	0.30
r(BS/NucPi)	-0.37	0.32	r(SLGM/NucPi)	0.28	0.32
			r(SPD/MtHD)	0.51	0.02*
			r(SPD/MtPi)	0.53	0.02*
			r(SPD/NucHD)	0.07	0.68
			r(SPD/NucPi)	0.06	0.74
			r(S2080/MtHD)	0.46	0.04*
			r(S2080/MtPi)	0.26	0.25
			r(S2080/NucHD)	0.12	0.56
			r(S2080/NucPi)	0.09	0.62

The range of Φ_{ST} pairwise comparisons for the *MT-CYB* gene was 0.002-0.574 (Table S4 in the Supporting Information of this chapter). The majority of comparisons among populations showed significant differences, even after Bonferroni's correction ($p < 0.0001$). These results were confirmed by hierarchical AMOVA (Table 2). Hence, significant levels of population differentiation were found for both loci when we analysed the entire data set as well as when basins were analysed as independent groups. Significant levels of genetic structure among river within basins and within rivers for both loci were also supported. SAMOVA analyses were congruent with AMOVA outcomes.

Table 2. Genetic hierarchical population structure of *Squalius valentinus* populations based on *MT-CYB* y *RAG1*. * means p-value significant in SAMOVA.

Structure tested	MT-CYB			RAG1		
	% Variance	Φ statistics	p-value	% Variance	Φ statistics	p-value
Global diversity						
(one gene pool)						
Among basins	39.4	$\Phi_{ST} = 0.394$	<0.0001	6.9	$\Phi_{ST} = 0.069$	<0.0001
Within basins	60.5			93.13		
Independent basins						
(nine gene pools)						
Among basins	6.6	$\Phi_{CT} = -0.066$	0.66599*	4.1	$\Phi_{CT} = 0.042$	0.02542
Among rivers within basins	45.6	$\Phi_{SC} = 0.428$	<0.0001	3.8	$\Phi_{SC} = 0.033$	0.00391
Among rivers	61.0	$\Phi_{ST} = 0.390$	<0.0001	92.1	$\Phi_{ST} = 0.073$	<0.0001

2. Population size changes and migration rate

Based on *MT-CYB*, confirmation of population expansion was based on neutrality tests. For the complete data set, Fu's F_S and Tajima's D gave significant negative values due to large number of low frequency haplotypes (Table S5 in the Supporting Information of this chapter). Mismatch distribution including all populations was unimodal and was adjusted to the distribution inferred by the growth-decline population model (Fig. S1A in the Supporting information of this chapter). Happerding's raggedness index (r) corroborated the fit of the observed and expected distributions (Table S5 in the Supporting Information of this chapter). When basins were considered independently, negative values of Fu's F_S and Tajima's D , as well as unimodal mismatch distribution, were found in most of the basins (Fig. S2A in the Supporting Information of this chapter). Nevertheless, levels were only significant in those with the highest mitochondrial genetic diversity values (Table S5 in the Supporting Information of this chapter). The largest of the northern basins exhibited multimodal distribution (Fig. S2A in the Supporting Information of this chapter). When mismatch was performed independently on its tributaries, those from middle and upper reaches (with a pluvio-nival precipitation regime) showed unimodal distribution, and those from downstream (alluvial plains) were multimodal (Fig. S2B in the Supporting Information of this chapter).

Based on τ values, a generation time of 2 years and an evolutionary rate of 0.4% per million years per lineage produced a mutation rate (μ) of 8.68×10^{-6} per nucleotide per year for the *MT-CYB* gene. Using these data, we estimated time since expansion (t) in mutational units for

the complete data set and for those independent basins showing significant population expansion from 190,000 to 50,000 years ago (Table S5 in the Supporting Information of this chapter). The EBSP analysis of both genes showed an exponential population expansion in all basins approximately 50,000 years ago following a period of consistent population growth (Fig. 1SB in the Supporting Information of this chapter). Similar patterns of exponential or more gradual population expansion were further corroborated by EBSP analyses in those basins with significant neutrality tests (Fig. S3 in the Supporting Information of this chapter). Estimates of historical migration rates among basins based on the *MT-CYB* gene showed highly symmetrical magnitude and direction in most basins (Table S6 in the Supporting Information of this chapter).

3. Species distribution modelling and landscape genetic approaches

With respect to the inference of habitat suitability, estimated models showed small changes, over time, in *S. valentinus* habitat suitability, although some areas showed similar high values (>0.7) for all historical periods (Fig. 4). Accuracy of all models was high ($AUC > 0.9$), providing acceptable predictive power (Araujo *et al.*, 2005). The two thresholds chosen as dividing line between suitable and unsuitable habitats recovered some overlapping areas across all periods where habitat is suitable for the species (Table 3; Fig. 5).

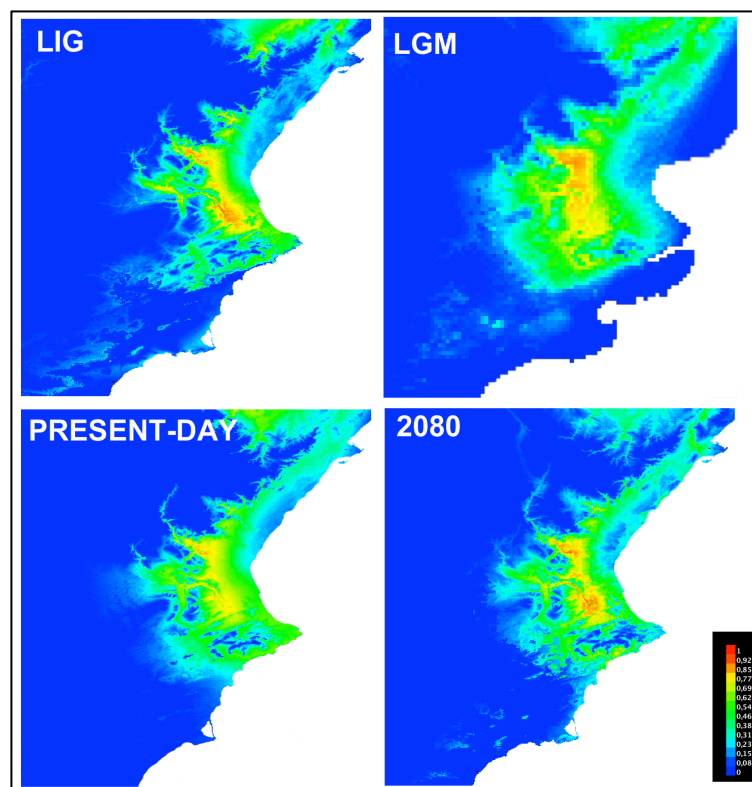


Fig. 4. *Squalius valentinus* hábitat suitability modelling. Inferred habitat suitability. Colours indicate the probability of occurrence (see legend).

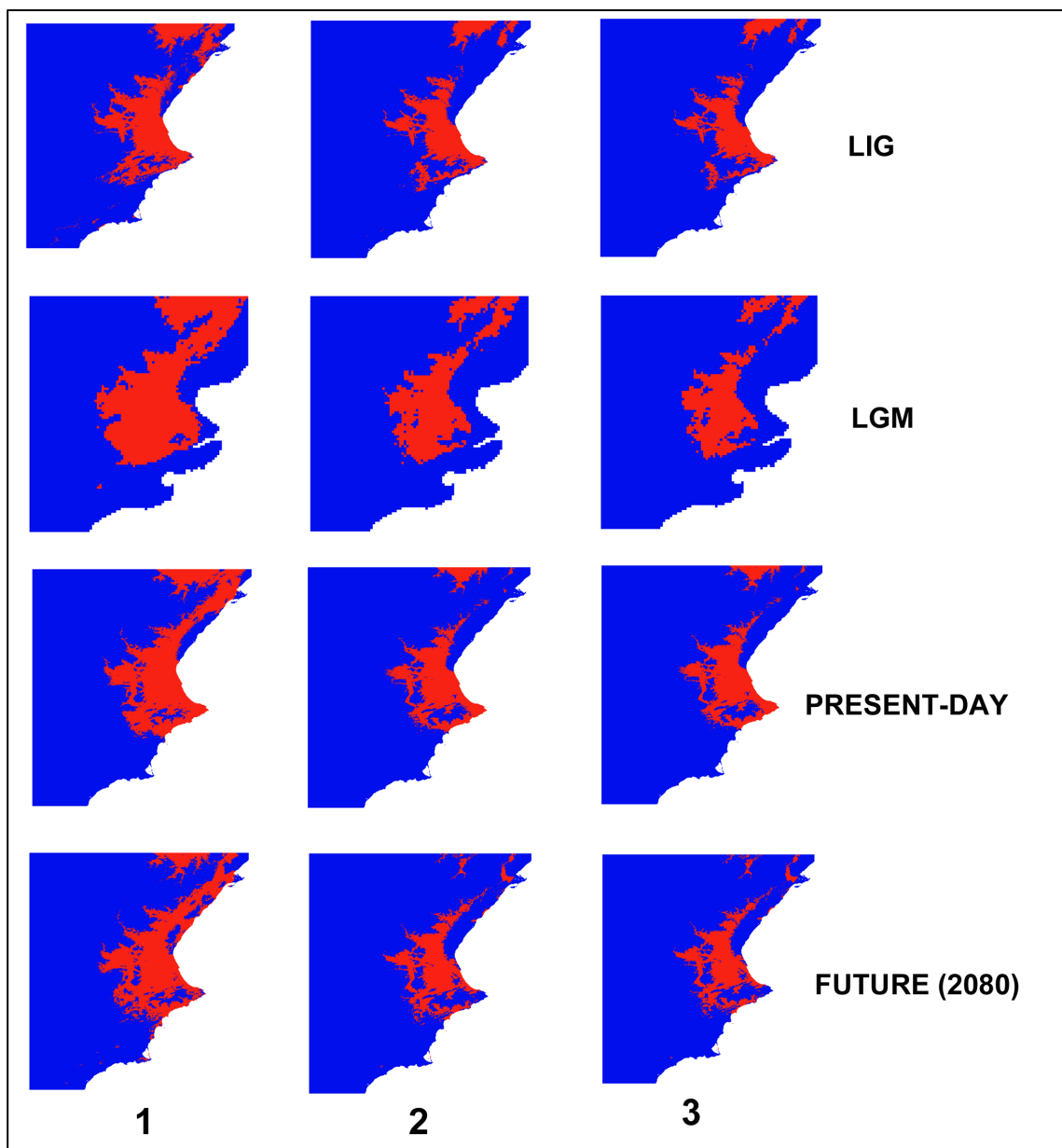


Fig. 5. (A) Binary suitability/unsuitability maps for all historical periods according to two threshold levels. (B) Habitat suitability overlapping for the four historical periods according to the two thresholds chosen.

Table 3. Threshold values for binary models distinguishing between suitable and unsuitable habitats in all historical period analysed. Minimum, minimum training presence logistic threshold; 10 percentile, 10 percentile training presence logistic threshold.

	Last Interglacial	Last Glacial Maximum	Present day	2080
Minimum	0.21	0.22	0.23	0.21
10 percentile	0.31	0.35	0.36	0.35

According to the regularized training gain obtained by the jackknife test after eliminating correlated variables (see Appendix S1 in the Supporting Information of this chapter), for current conditions modelling, the two variables with the highest gain when considered independently were BIO6 *mean temperature of coldest month* and BIO7 *temperature annual range*; thus, it provides the most useful information. The climatic variable that decreased the gain to the greatest extent when omitted was also BIO6. Hence, they represent the largest amount of information that was not given by any other variable. In the present-day period, BIO6 was negatively related to habitat suitability according to the variable response curve obtained in MAXENT for this variable (data not showed). In the LIG model, the variable with the most useful information was BIO7 *temperature annual range*, which was also negatively related to habitat suitability. For the LGM, BIO7 *temperature annual range* and BIO18 *precipitation of the warmest quarter* were the variables with the highest gain considered independently and decreasing the gain to the greatest extent, respectively. Higher habitat suitability values were related to medium values of both variables. Predictions for the future (year 2080) revealed BIO14 *precipitation of driest month* and BIO19 *precipitation of coldest quarter* to be the variables that contribute most to the model.

We also found a significant positive correlation between habitat suitability with mitochondrial, but not with nuclear, genetic diversity, measured as haplotype and nucleotide diversity (Table 1), as well as with genetic structure, measured as F_{ST} values. In the latter case, regression analyses performed in GESTE showed the model including present-day habitat suitability to have the highest Bayesian probability and, hence, the model better explaining genetic structure of *S. valentinus* populations (Table 4). GESTE analyses also demonstrated that the bioclimatic variables with the strongest effect on the genetic structure of this species were always related to seasonality and “extreme” climatic conditions regardless of the combination of variables used in the models and the historical period considered (Table 5, S7 and S8 in the Supporting Information of this chapter).

Table 4. Posterior probability values rendered from testing the effect of habitat suitability (S) on genetic structure of *Squalius valentinus* populations using GESTE software. Posterior probabilities of regression models including the variable habitat suitability and best regression model explaining genetic structure of this species.

	Sum of posterior probability	Model with the highest posterior probability (p-value)
Last Interglacial	0.18	-
Last Glacial Maximum	0.25	-
Present-day period	0.76	2.4 – 0.615 S (p=0.353)

Table 5. Posterior probabilities of regression models including bioclimatic variables (in bold) that best explain genetic structure of *Squalius valentinus* populations for each data set. Models were obtained using GESTE software. * means no effect of bioclimatic variables

Historical period	Regression Model	Bayesian probability
LIG	TEMPERATURE: -3.63 – 1.83 BIO11 + 2.93 BIO7 – 2.65 BIO6	0.063
	PRECIPITATION: 2.32 – 0.512 BIO15	0.104
	SEASONALITY: 2.3 + 0.554 BIO7	0.374
LGM	“EXTREME” CLIMATIC CONDITIONS: 2.29 – 0.618 BIO6	0.529
	TEMPERATURE: -0.911 + 2.34 BIO9	0.032
	PRECIPITATION: 0.257 – 4.02 BIO18	0.087
	SEASONALITY: 2.27 + 0.522 BIO14	0.335
	“EXTREME” CLIMATIC CONDITIONS: 2.28*	0.294
Preset day	TEMPERATURE: 0.403 + 3.20 BIO10 + 2.53 BIO3 + 1.69 BIO2 + 2.97 BIO1	0.101
	PRECIPITATION: -2.37 – 1.58 BIO16 – 1.78 BIO15 – 1.55 BIO13	0.042
	SEASONALITY: 2.3 + 0.529 BIO4	0.316
	“EXTREME” CLIMATIC CONDITIONS: 2.25*	0.276

DISCUSSION

Results of the study demonstrate the important contribution that basin size and geomorphology have exerted on the genetic structure and diversity of Mediterranean-type freshwater species, as well as the value of stable aquatic systems, such as those fed by springs, in coping with the severe fluctuating conditions imposed by the Mediterranean climate.

1. Mediterranean seasonality effects on genetic diversity, structure and demography of freshwater fishes

Based on constraints imposed by the Mediterranean climate, we would expect a well-supported genetic structure, induced by fragmentation events, in freshwater fish populations inhabiting Mediterranean-type river systems. Our results supported a significant genetic structure in *S. valentinus* for mitochondrial and nuclear genes (significant Φ_{ST} , AMOVA and SAMOVA). This genetic structure among basins, which has also been described for other Iberian freshwater fish species (Doadrio and Carmona, 2004; Almada and Sousa-Santos, 2010; Lopes-Cuhna *et al.*, 2012), can be linked to historical palaeogeographical factors related to the Plio-Pleistocene formation of the current exohereic Iberian basins (Calvo *et al.*, 1993) and to a lesser extent to the Mediterranean climate.

Among rivers within basins and within individual rivers, we found significant genetic structure that can be associated with differential effects of Mediterranean seasonality on areas within the basin. This is especially evident in the largest of the northern basins, which display distinct fluvial regimes in which tributaries or stretches within a tributary are subjected to cyclic flooding and drought periods differing in magnitude and frequency. In this case study, species genetic diversity and genetic structure were higher in downstream tributaries within the largest basin, as revealed by the differentiated haplogroup with unique haplotypes not found in the upper tributaries or in the main course (Fig. 2A). This is in agreement with the typical geomorphological configuration found in the Mediterranean-type basins, which usually present downstream alluvial plains in which the effect of drought is more severe than in upper reaches having a steeper river profile. This highlights the key role of river geomorphology in shaping the genetic diversity and structure of freshwater populations (Hughes, 2007; Huey *et al.*, 2008). If populations isolated by seasonality do not suffer strong population decline leading to extinction, they may reach demographic stability, as we found for *S. valentinus* populations inhabiting downstream areas. The maintenance of stable populations has allowed, over time, development of high levels of genetic diversity and structure, especially in regions under the influence of seasonal weather patterns (Sousa-Santos *et al.*, 2014). Restriction of gene flow and population differentiation is especially marked in Mediterranean regions of sem-arid and arid climate, in which only unpredictable floods restore connectivity among streams and waterholes (Faulks *et al.*, 2010). In steeper upper reaches, fluvial connectivity is higher, and evolutionary processes such as migration and population expansion prevail over genetic drift and bottlenecks, preventing the formation of a clear genetic structure and diminishing the genetic diversity in these areas, as has been demonstrated in other Mediterranean-type freshwater fish species (Faulks *et al.*, 2010; Henriques *et al.*, 2010).

The influence of the Mediterranean climate was expected to be stronger in the smaller southern basins and in the smaller tributaries of the larger northern basins, as typical Mediterranean alternating flow regimes may be extreme in small basins, especially in intermittent or temporary stretches. Extreme seasonality exacerbates consequences of density-dependent relationships, as occurs when fish are confined in ponds during summer drought, increasing biotic interactions (Gasith and Resh, 1999; Matthews and Marsh-Matthews, 2003) and affecting survival and recruitment of populations of each generation (Poizat and Crivelli, 1997). Over time, this may modify the demographic structure and genetics of populations of Mediterranean-type species such *S. valentinus*, as environmental conditions become more critical, and processes such bottlenecks are more common (Magalhães *et al.*, 2003).

For *Squalius* species, drainage area size has shown a strong correlation with genetic diversity (Almada and Sousa-Santos, 2010). This was found in the present study for nuclear, but not for mitochondrial, data (Table 1). Unexpectedly, we found low mitochondrial, but not nuclear, genetic diversity values in a single northern basin (Fig. 3). Genetic drift following a bottleneck event (positive and significant F_s ; Fu's F_s ; Rand, 1996) may be the primary source of the lower mitochondrial genetic diversity in this basin. Similar mitochondrial bottlenecks have been found in other fish species (Johannesson and André, 2006). Contrary to the concept that larger basins express higher genetic diversity due to large historical populations, we found the highest values of mitochondrial genetic diversity in small basins containing deep ponds and/or permanent springs, probably associated with habitat suitability and habitat stability. The lowest genetic diversity values were found in the smaller southern intermittent basins. These low values may be the result of disproportional vulnerability to intermittent extreme flooding. The main effects of erratic flood/droughts cycles are the reduction in effective population size (bottlenecks) and, consequently, the reduction in genetic diversity (Huey *et al.*, 2006; Hughes and Hyllier, 2006). This lower diversity found in the intermittent southern rivers and streams may also be associated with bottlenecks resulting from founder effects (Roman, 2006), as these basins constitute the southern limits of *S. valentinus* distribution.

Despite our finding of significant genetic structure in *S. valentinus* that may be associated with Mediterranean seasonality, the inferred haplotype networks did not exhibit the typical shape found for Mediterranean-type species, that of abundant and low frequency haplotypes grouped into well-differentiated haplogroups (e.g. Snoj *et al.*, 2009; Sousa-Santos *et al.*, 2014). In our study, networks presented different haplogroups or unique haplotypes in different basins, but all were connected to a central haplotype, conferring a stellate shape to the network (Fig. 2A,B). These stellate networks are characteristic of populations in expansion, as we found in *S.*

valentinus, and, as shown in other studies, sometimes of populations that do not have significant genetic structure (Dohms and Burg, 2013). Most Spanish Levantine rivers run for short distances before flowing into the Mediterranean Sea. During annual rainfall periods, extreme floods are common, allowing fluvial connectivity restoration and mixture of populations and providing a contemporary factor to explain the star-shaped haplotype networks.

A further historical factor responsible for the observed stellate networks is the attenuating influence of Pleistocene glaciations on the genetic structure of *S. valentinus* populations. Although ice sheets did not directly cover most of the Spanish Levantine area during glaciations, glaciers had an important impact on fluvial dynamics (Hughes *et al.*, 2006; Hughes and Woodward, 2008), altering genetic and demographic structure rather than depleting populations. The star-shaped structure of networks and the sharing of the most common haplotype and allele by all populations point to a single glacial refuge for *S. valentinus* that remained over time owing to the high connectivity among basins, as supported by the high historical migration rates revealed by MIGRATE-N analysis. This connectivity among basins was inferred in our bathymetric map reconstruction (Fig. 1) and was the effect of downstream connections resulting from a drastic drop in sea level during the LGM, approximately 21,000-18,000 years ago (Waelbrock *et al.*, 2002; Miller *et al.*, 2005). Present-day Europe experienced two glacial periods during this interval: Riss (c. 200,000-140,000) and Würm (c. 120,000-10,000), which likely promoted the extension of the hydrological network following sea-level retraction and allowed the expansion of *S. valentinus* populations as supported by the neutrality tests for the complete data set and for some of the independent basins (Table S6 in the Supporting Information of this chapter).

The continental shelf of the Spanish Levantine region is wide and shallow, especially in the northernmost area (Casas-Sainz and de Vicente, 2009), favouring potential extension of downstream riverine areas during glacial episodes and hence mixing populations. Inferred basins, the limits of which lie between current (0 m) and LGM (-135 m) shorelines, in the northernmost areas were not well defined, suggesting that, during glacial periods, the region may have been a shallow continually flooded area in times of heavy rainfall. Similar broad continental shelf and hypothetical lowland connections among drainages during ancient periods of low sea levels, and their consequences for freshwater organisms, have been described in other Mediterranean-type regions (Swartz *et al.*, 2007; Unmack *et al.*, 2012).

In contrast, bathymetric reconstruction showed that most of the southern smaller basins covered longer stretches during glacial periods, but without downstream connections (Fig. 1). For this reason, relationships among some of these southern basins and population expansion may

be better explained by river piracy or river interconnections during strong floods (e.g. Burridge *et al.*, 2007; Swartz *et al.*, 2007). No palaeogeographical evidence of river capture is known in the Spanish Levantine region; nevertheless, some of these rivers are enclosed within a marshy area that probably accounts for the sharing of haplotypes and alleles. In the largest northern basin with two fluvial regimes, the Júcar, although EBSF supported population expansion, we found a negative non-significant value of F_u 's F_s . Furthermore, mismatch analysis did not fit with an expansion model and suggested that population expansion occurred in the middle-upper reaches but not in downstream area, which seem to have experienced long-term demographic stability.

2. Species distribution modelling and landscape genetics approaches

Distribution modelling showed changes associated with climatic variation in inferred potential distribution of *S. valentinus* over time conditioned to habitat suitability. The inferred distribution was more extensive during the earlier analysed period, the LIG, when the climate conditions in Iberian Levantine basins, as well as in other central Iberian basins, seemed more suitable for *S. valentinus* populations than was the case in later periods. After this time, models showed a slight ongoing reduction in potential distribution that is especially evident in future projections (Fig. 4), suggesting that the climate is becoming less favourable for *S. valentinus*, a common finding in Mediterranean-climate ecosystems (Bálint *et al.*, 2011).

Some areas of the distribution range of *S. valentinus* presented similar high values of habitat suitability for all analysed periods, including the future, suggesting suitable climatic conditions for *S. valentinus* in those regions (Figs. 4 and 5). These areas may be considered the zones of stability for the species, defined as places that can potentially be occupied by a population or a species irrespective of the time period (Carnaval *et al.*, 2009). The higher habitat suitability values for all time periods were found chiefly near the Albufera Lagoon and small tributaries of the downstream areas of the larger northern basin drainages (Figs. 1 and 4). The mentioned areas of stability are characterized by springs typical of Spanish Mediterranean littoral basins that supply permanent streams with water from subterranean aquifers as well as from surface run-off of surrounding mountains (Cantoral-Uriza and Aboal, 2010), thereby buffering the seasonal water-level fluctuations associated with Mediterranean climates. These areas of stability in Mediterranean-type systems provide a refuge of stable environmental conditions for freshwater fish (Magooulick and Kobza, 2003). Deep ponds or waterholes as refugia are essential in areas in which Mediterranean seasonality is extreme, such as in semi-arid and arid regions, to ensure the survival of freshwater populations until fluvial connectivity is restored (Balcome *et al.*, 2007; Huey *et al.*, 2008; Faulks *et al.*, 2010). Studies have demonstrated that stable environments contribute to developing and maintaining high genetic diversity over time (Fauvelot *et al.*, 2003), as we

found in these areas of stable climatic conditions (Fig. 3). These stable climatic areas are also found in basins that have undergone significant population expansion (northern basins and southern basins with springs in our study) or demographic stability (tributaries downstream of Júcar in our study), favouring larger effective population sizes over time.

The bioclimatic variables contributing most to all modelled periods indicated the interplay between temperature during the warmest and coldest season and precipitation during the driest season in determining the habitat suitability for *S. valentinus*. In the LGM, one of the variables contributing the most suggests the importance of precipitation, in one of the driest areas of the Iberian territory, to supporting freshwater communities during the warmest season. In addition, in the LGM, and also in LIG, periods, the temperature annual range was a notable factor for the inferred potential distribution of the species conditioned by environmental conditions; this variable was negatively correlated to habitat suitability in both periods and demonstrated that stable environmental conditions, in this case temperature, were crucial to maintaining *S. valentinus* populations along time, especially important during the LGM when climatic conditions become less favourable for organisms. Species distribution projections for the future also highlight the importance of precipitation to ensure the preservation of the species in the Spanish Levantine region, which may be extended to include Mediterranean freshwater ecosystems in general, where ongoing aridification is a common trend (Seager *et al.*, 2014).

According to habitat suitability, our data support a high structuring of populations associated with the present-day model (Table 4) and demonstrate that lower values of habitat suitability favour higher level of genetic structure. Models including probability of occurrence in LGM and future predictions showed also higher Bayesian probabilities relative to the rest of models (0.144 and 0.150 respectively), which not reached probability values above 0.08. Therefore, we support a strong negative correlation between habitat suitability and genetic structure (fragmented populations).

With respect to the bioclimatic variables having the strongest effect on genetic structure of *S. valentinus* populations, higher values of Bayesian probability were found in models that included variables associated with seasonality in all historical analysed periods and to “extreme” climatic conditions in the LGM and the future predictions. For LIG, LGM and the present-day period, seasonality variables were positively correlated with genetic structure, while the correlation between this latter factor and “extreme” climatic variables was negative. These findings support the influence of Mediterranean seasonality on the genetic structure of *S. valentinus* since a long time ago. At Mediterranean latitudes, periods of high temperature and lower levels of precipitation are related to summer droughts, a natural disturbance that can

promote population structuring (Magoulick and Kobza, 2003), as consequence of fragmentation events and their evolution-associated processes, which have been previously discussed. The association of seasonality and “extreme” climatic variables with genetic structure suggests a future trend to increase population structuring to the extent that precipitation decrease (Phillandras *et al.*, 2011; Seager *et al.*, 2014).

3. Conservation of Mediterranean-type freshwater fish

The large increases in human population of Mediterranean regions in recent decades, with its associated activities such as agriculture, tourism and industry, have exacerbated depletion of water resources (Underwood *et al.*, 2009). Mediterranean floodplain areas, such as those found in downstream stretches in the northern basins analysed here, constitute one of the most agriculturally productive regions of the world and are especially sensitive to hydrological alterations (Tockner *et al.*, 2008). For this reason, water abstraction for agriculture irrigation represents a critical threat to Mediterranean-type freshwater ecosystems (Underwood *et al.*, 2009 and references herein). Precipitation is the limiting factor of the preservation of aquatic organisms in Mediterranean-type regions and is especially critical as longer periods of droughts and greater interannual weather variability are considered to be increasing trends as the climate warms (Giorgi and Lionello, 2008; Orlowsky and Seneviratne, 2012). Human exploitation of water resources alters the natural flood/drought cycle to which Mediterranean-type species are adapted and consequently increases the risk of extinction of these species, including *S. valentinus*.

CONCLUSIONS

The population history of Mediterranean freshwater fish species such as *S. valentinus* during the Plio-Pleistocene must be viewed as an interaction among climate fluctuations and adaptability to new ecological opportunities influenced by fluctuating intra- and interannual seasonality. The interplay between factors such as the Mediterranean climate and the Pleistocene glaciations, along with the end of the current Iberian hydrographic configuration, is responsible for the phylogeographic and demographic structure of *S. valentinus* populations. Our findings provide evidence for the important role of the stable environmental conditions in habitat refuge provided by small springs for maintaining genetically diverse populations, such as occurs with *S. valentinus* in the Spanish Levantine region. The populations from stable habitats retain high values of haplotype and nucleotide diversity. Fluctuating climatic conditions occurring in Plio-Pleistocene periods associated with Mediterranean climate favoured the development of two demographic models in the Spanish Levantine region for *S. valentinus*. Populations in small and

extremely seasonal streams are more susceptible to influence by extreme droughts/floods and more likely to suffer decline in effective population size. Basins with permanent freshwater areas allowed population expansion or maintenance of a consistent effective population size. Although these latter areas are also subjected to flooding and drought, the impact is not usually so severe. Our results offer the opportunity to extrapolate findings to other species that inhabit seasonal Mediterranean river systems.

Our study shows the utility of species distribution models and landscape genetics in evolutionary studies. This methodology allows linking of environmental and climatic forces to the potential distribution of species, conditioned to habitat suitability, in various time periods and to the genetic diversity and structure of populations. We applied modelling of potential species distribution relative to climate variables and found, for *S. valentinus*, that rainfall during the warmest periods and narrow temperature annual range to be the factors having greatest impact on habitat suitability in all the analysed periods. This provides evidence of the importance of habitat suitability, such as is provided by thermal springs of karstic origin that act as refuges for freshwater species in Mediterranean fluctuating systems. These springs are common in the Mediterranean region and are mainly supplied by subterranean aquifers near coastal areas. However, they are overexploited, mainly for irrigation, due to being located in one of the most productive agricultural regions in the world, as well as one that has seen high impact of tourism. This study also demonstrated the strong correlation between habitat suitability and genetic structure and the importance of climatic variables related to seasonality in driving the genetic structure of populations of Mediterranean species.

Acknowledgements

We thank P. Garzón, F. Alda, J.L. González, P. Risueño and E. Aparicio for field assistance. L. Alcaraz, I. Polo and I. Sagar helped with laboratory procedures. F. Morcillo gave us helpful contribution for bathymetric analyses. The Lucidus Consultancy helped us with English editing of the manuscript. We are also grateful to one nonanonymous reviewer (J. robalo) and two anonymous reviewers who greatly improved the final version of the manuscript with their suggestions and constructive criticisms. This study had financially sponsored by the Spanish CGL2010-15231BOS project and by the Valencian Government (Conselleria de Medio Ambiente de la Comunidad Valenciana).

DATA ACCESSIBILITY

- Genbank Accesion numbers: *MT-CYB*: KR871719-KR871767; *RAG1*: KR871768-KR871795 (Table S1 in the Supporting Information of this chapter).
- MAXENT presence-data coordinates for *Squalius valentinus* (Table S2 in the Supporting Information of this chapter).
- Environmental variables analysed in GESTE (Appendix S1 in the Supporting Information of this

chapter).

-Genetic data input file used in GESTE (Appendix S2 in the Supporting Information of this chapter).

-Dryad entry for sequence alignments (<http://dx.doi.org/10.5061/dryad.sh5530>).

REFERENCES

- Almada VC, Sousa-Santos C. 2010. Comparisons of the genetic structure of *Squalius* populations (Teleostei, Cyprinidae) from rivers with contrasting histories, drainage area and climatic conditions based on two molecular markers. *Molecular Phylogenetics and Evolution*, 57: 924-931.
- Araujo MB, Pearson RG, Thuiller W, Erhard M. 2005. Validation of species-climate impact models under climate change. *Global Change Biology*, 11: 1504-1513.
- Balcombe SR, Bunn SE, Arthington AH, Fawcett JH, McKenzie-Smith FJ, Wright A. 2007. Fish larvae, growth and biomass relationships in an Australian arid zone river: links between floodplains and waterholes. *Freshwater Biology*, 52: 2385-2398.
- Bálint M, Domisch S, Engelhardt CHM, Haase P, Lehrian S, Sauer J, Theissinger K, Pauls SU, Nowak C. 2011. Cryptic biodiversity loss linked to global climate change. *Nature Climate Change*, 1: 313-318.
- Beerli P. 2009. How to use MIGRATE or why are Markov chain Monte Carlo programs difficult to use? In: Bertorelle G, Bruford MW, Hauffe HC, Rizzoli A y Vernesi C (eds.). *Population Genetics for Animal Conservation*. Cambridge University Press. Cambridge. UK. Pp 42-76.
- Benito G, Thorndycraft VR, Rico M, Sánchez-Moya Y, Sopeña A. 2008. Paleoflood and floodplain records from Spain: evidence for long-term climate variability and environmental changes. *Geomorphology*, 101: 68-77.
- Blondel J y Aronson J. 1999. *Biology and Wildlife of the Mediterranean Region*. Oxford University Press.
- Boix D, García-Berthou E, Gascón S, Benejam L, Tornés E, Sala J, Benito J, Munné A, Solá C, Sabater S. 2010. Response of community structure to sustained drought in Mediterranean rivers. *Journal of Hydrology*, 383: 135-146.
- Bond NR, Lake PS, Arthington AH. 2008. The impacts of drought on freshwater ecosystems: an Australian perspective. *Hydrobiologia*, 600: 3-16.
- Bruent T, Phillipe H, Bryant D. 2006. A quick and a robust statistical test to detect recombination. *Genetics*, 172: 2665-2681.
- Burridge CP, Craw D, Waters JM. 2007. An empirical test of freshwater vicariance via river capture. *Molecular Ecology*, 16: 1883-1895.
- Calvo JP, Daams R, Morales J, López-Martínez N, Agustí J, Anandon P, Armenteros I, Cabrera L, Civis J, Corrochano A, Díaz-Molina M, Elizaga E, Hoyos M, Martín-Suarez E, Martínez J, Moissenet E, Muñoz A, Pérez-García A, Pérez-González A, Portero JM, Robles F, Santisteban C, Torres T, Van der Meulen AJ, Vera JA, Mein P. 1993. Un to date Spanish continental synthesis and paleoclimatic interpretation. *Revista de la Sociedad Geológica de España*, 6: 29-40.
- Cantoral-Uriza EA, Aboal M. 2010. Comunidades algales bentónicas de algunas Fuentes del marjal de Oliva-Pego, Comunidad Valenciana. España. *Limnetica*, 29: 287-300.
- Carnaval AC, Hickerson MJ, Haddad CFB, Rodrigues MT, Moritz C. 2009. Stability predicts genetic diversity in the Brazilian Atlantic Forest hotspot. *Science*, 5915: 785-789.
- Casas-Sainz AM, de Vicente G. 2009. On the tectonic origin of Iberian topography. *Tectonophysics*, 474: 214-235.

- Davey AJH, Kelley DJ. 2007. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, 144: 2001-2014.
- Dekar MP, Magoulick DD. 2007. Factors affecting fish assemblage structure during seasonal stream drying. *Ecology of Freshwater Fish*, 16: 335-342.
- Doadrio I, Carmona JA. 2003. Testing freshwater Lago Mare dispersal theory on the phylogeny relationships of Iberian cyprinids genera *Chondrostoma* and *Squalius* (Cypriniformes, Cyprinidae). *Graellsia*, 59: 457-473.
- Doadrio I, Carmona JA. 2006. Phylogentic overview of the genus *Squalius* (Actinopterygii, Cyprinidae) in the Iberian Peninsula, with description of two new species. *Cybium*, 30: 199-214.
- Doadrio I, Perea S, Alcaraz L, Hernández N. 2009. Molecular phylogeny and biogeography of the Cuban genus *Girardinus* Poey, 1854 and relationships within the tribe Girardinini (Actinopterygii, Cyprinidae). *Molecular Phylogenetic and Evolution*, 50: 16-30.
- Doadrio I, Perea S, Garzón-Heydt P, González JL. 2011. *Ictiofauna continental española. Bases para su seguimiento*. DG Medio Natural y Política Forestal. MARM. Madrid.
- Dohms KM, Burg TM. 2013. Molecular markers reveal limited population genetic structure in a North American corvid, Clark's nutcracker (*Nucifraga columbiana*). *PLoS One*, 8: e79621.
- Domínguez-Domínguez O, Vila M, Pérez-Rodríguez R, Remón N, Doadrio I. 2011. Complex evolutionary history of the Mexican stoneroller *Campostoma ornatum* Girard, 1856 (Actinopterygii, Cyprinidae). *BMC Evolutionary Biology*, 11: 113.
- Drummond A, Rambaut A. 2007. BEAST: Bayesian Evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, 7: 214.
- Drummond A, Rambaut A, Shapiro B. 2005. Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular Biology and Evolution*, 22: 1185-1192.
- Dupaunloup I, Schneider S, Excoffier L. 2002. A simulated annealing approach to define the genetic structure of populations. *Molecular Ecology*, 11: 2571-2581.
- Elith J, Graham CH, Anderson RP, Dudík M, Simon F, Guisan A, Hijmans R, Huetteman F. 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, 29: 129-151.
- Excoffier L y Lischer HEL. 2010. Arlequin suite ver. 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, 10: 564-567.
- Faulks LK, Gilligan DM y Beheregay LB. 2010. The role of anthropogenic vs natural in-stream structure in determining connectivity and genetic diversity in an endangered freshwater fish, Macquarie perch (*Macquaria australasica*). *Evolutionary applications*, 4: 589-601.
- Fauquette S, Suc J-P, Guiot J, Diniz F, Feddi N, Zheng Z, Bessais E y Drivaliari A. 1999. Climate and biomes in the west Mediterranean area during the Pliocene. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 152:15-36.
- Fauvelot C, Bernardi G y Planes S. 2003. Reductions in the mitochondrial DNA diversity of coral reef fish provide evidence of population bottleneck resulting from Holocene sea-level change. *Evolution*, 57: 1571-1583.
- Foll M y Gaggiotti OE. 2006. Identifying the environmental factors that determine the genetic structure of populations. *Genetic*, 174: 875-891.
- Frankham R. 1996. Relationships of genetic variation to population size in wildlife. *Conservation Biology*, 10: 1500-1508.
- Fu XY. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, 147: 915-925.
- Fuentes-Utrilla P, Venturas M, Hollingsworth PM, Squirrel J, Collada C, Stone GN y Gil L. 2014. Extending glacial refugia for European tree: genetic markers show that Iberian populations of white elm are native relicts and not introductions. *Heredity*, 112: 105-113.
- Gaggiotti OE, Bekkevold D, Jorgensen HBH, Foll M, Carvalho GR, Andre C y Ruzzante DE. 2009. Disentangling the

- effects of evolutionary, demographic, and environmental factors influencing genetic structure of natural populations: Atlantic herring as a case study. *Evolution*, 63: 2939-2951.
- Gante HF, Micael J, Oliva-Paterna FJ, Doadrio I, Dowling TE y Alves MJ. 2009. Diversification within glacial refugia: tempo and mode of evolution of the polytypic fish *Barbus sclateri*. *Molecular Ecology*, 18: 3240-3255.
- Gasith A y Resh VH. 1999. Streams in Mediterranean climate regions: abiotic influences and biotic responses to predictable seasonal events. *Annual Review of Ecology and Systematics*, 30: 51-81.
- Giorgio F y Lionello P. 2008. Climate change projections for the Mediterranean region. *Global and Planetary Change*, 63: 90-104.
- Gómez A y Lunt HD. 2007. Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula. In: Weiss S y Ferrand N (eds.). *Phylogeography of Southern European Refugia*. Springer. Dordrecht. The Netherlands. Pp 155-188.
- Graham MH. 2003. Confronting multicollinearity in ecological multiple regression. *Ecology*, 84: 2809-2815.
- Haywood AM, Dowsett HJ, Valdes PJ, Lunt DJ, Francis JE y Sellwood BW. 2009. Introduction. Pliocene climate, processes and problems. *Philosophical Transaction of Royal Society Series A*, 367: 3-17.
- Henriques R, Sousa V, Coelho MM. 2010. Migration patterns counteract seasonal isolation of *Squalius torgalensis*, a critically endangered freshwater fish inhabiting a typical Circum-Mediterranean small drainage. *Conservation Genetics*, 11: 1859-1870.
- Hewitt GM. 2004. Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transaction of Royal Society Series B, Biological Sciences*, 359: 183-195.
- Hijmans RJ, Cameron SE, Parra JL, Jones PG y Jarvis A. 2005. Very high resolution interpolated climate surfaces for the global land areas. *International Journal of Climatology*, 25: 1965-1978.
- Hudson DJ y Bryant D. 2006. Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, 23: 254-267.
- Huey JA, Hughes JM y Baker AM. 2006. Patterns of gene flow in two species of eel-tailed catfish, *Neosilurus hyrtii* and *Porochilus argenteus* (Siluriformes: Plotosidae), in western Queensland's dryland rivers. *Biological Journal of the Linnean Society*, 87: 457-467.
- Huey JA, Baker AM y Hughes JM. 2008. The effect of landscape processes upon gene flow and genetic diversity in an Australian freshwater fish, *Neosilurus hyrtii*. *Freshwater Biology*, 53: 1393-1408.
- Hughes JM. 2007. Constraints on recovery: using molecular methods to study connectivity of aquatic biota in rivers and streams. *Freshwater Biology*, 52: 616-631.
- Hughes JM, Hyllier MJ. 2006. Mitochondrial DNA and allozymes reveal high dispersal abilities and historical movement across drainage boundaries in two species of freshwater fishes from inland rivers in Queensland, Australia. *Journal of Fish Biology*, 68(B): 270-291.
- Hughes PD y Woodward JC. 2008. Timing of glaciation in the Mediterranean mountains during the last cold stage. *Journal of Quaternary Science*, 23: 575-588.
- Hughes PD, Woodward JC y Gibbard PL. 2006. Quaternary glacial history of Mediterranean mountains. *Progress in Physical Geography*, 30: 334-364.
- Humphries P, Baldwin DS. 2003. Drought and aquatic ecosystems: an introduction. *Freshwater Biology*, 48: 1141-1146.
- Jiménez-Moreno G, Suc J-P y Fauquette S. 2010. Miocene to Pliocene vegetation reconstruction and climate estimates from the Iberian Peninsula from pollen data. *Review of Paleobotany and Palynology*, 162: 403-415.
- Johannesson K y André C. 2006. Life on margins: genetic isolation and diversity loss in a peripheral marine ecosystem, the Baltic Sea. *Molecular Ecology*, 15: 2013-2029.
- Kottelat M y Freyhof J. 2007. *Handbook of the Freshwater Fish of Europe*. Cornol. Berlin.
- Lázaro R, Rodrigo FS, Gutiérrez L, Domingo F y Puigdefábregas J. 2001. Analysis of a thirty year rainfall record

- (1967-1997) from semi-arid SE Spain: a plant ecological perspective. *Journal of Arid Environments*, 48: 373-395.
- Librado P y Rozas J. 2009. DNASp v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, 25: 1451.
- Lopes-Cunha M, Aboim MA, Mesquita N, Alves MJ, Doadrio I, Coelho MM. 2012. Population genetic structure in the Iberian cyprinid fish *Iberochondrostoma lemmingii* (Steindachner, 1866): disentangling species fragmentation and colonization processes. *Biological Journal of the Linnean Society*, 105: 559-572.
- Lowe WH y Allendorf FW. 2010. What can genetics tell us about population connectivity? *Molecular Ecology*, 19: 3038-3051.
- Luikart G, England PR, Tallmon D, Jordan S, Taberlet P. 2003. The power and promise of population genomics: from genotyping to genome typing. *Nature Reviews*, 4: 981-994.
- Magalhães MF, Schlosser IJ, Collares-Pereira MJ. 2003. The role of life-history in the relationships between population dynamics and environmental variability in two Mediterranean stream fishes. *Journal of Fish Biology*, 63: 300-317.
- Magoulick DD, Kobza RM. 2003. The role of refugia for fishes during drought: a review and synthesis. *Freshwater Biology*, 48: 1186-1198.
- Matthews WJ, Marsh-Matthews E. 2003. Effects of drought on fish across axes of space, time and ecological complexity. *Freshwater Biology*, 48: 1232-1253.
- Miller KG, Kominz MA, Browning JV, Wright JD, Mountain GS, Katz ME, Sugarman PJ, Cramer BS, Christie-Blick N, Pekar SF. 2005. The Phanerozoic record of global sea-level change. *Science*, 310: 1293-1298.
- Nakicenovic N y Swart R. 2000. *IPCC Special Report: Emission Scenarios*. Cambridge University Press. Cambridge. England.
- Orlowsky B y Seneviratne SI. 2012. Global changes in extreme events: regional and seasonal dimension. *Climatic Change*, 110: 669-696.
- Ornelas-García CP, Domínguez-Domínguez O y Doadrio I. 2008. Evolutionary history of the fish genus *Astyanax* Baird y Girard (1854) (Actinopterygii, Cyprinidae) in Mesoamerica reveals multiple morphological homoplasies. *BMC Evolutionary Biology*, 8: 340.
- Pearson RG, Raxworthy CJ, Nakamura M y Peterson AT. 2007. Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *Journal of Biogeography*, 34: 102-117.
- Perea S, Böhme M, Zupančič P, Freyhof J, Šanda R, Ozuluğ M, Abdoli A y Doadrio I. 2010. Phylogenetic relationships and biogeographical patterns in Circum-Mediterranean subfamily Leuciscinae (Teleostei, Cyprinidae) inferred from both mitochondrial and nuclear data. *BMC Evolutionary Biology*, 10: 265.
- Phillandras CM, Nastos PT, Kapsomenakis J, Douvis KC, Tselioudis G y Zerefos CS. 2011. Long term precipitation trends and variability within the Mediterranean region. *Natural Hazards and Earth Systems Sciences*, 11: 3235-3250.
- Phillips SJ, Anderson RP y Schapire RE. 2006. Maximum entropy modelling of species geographic distributions. *Ecological Modelling*, 190: 231-259.
- Poizat G y Crivelli AJ. 1997. Use of seasonally flooded marshes by fish in a Mediterranean wetland: timing and demographic consequences. *Journal of Fish Biology*, 51: 106-119.
- Provan J y Bennet KD. 2008. Phylogeographic insights into cryptic glacial refugia. *Trends in Ecology and Evolution*, 23: 564-571.
- Rambaut A y Drummond A. 2007. Tracer v.1.4. Available at: <http://beast.bio.ed.ac.uk/Tracer>.
- Ramos-Onsins SE y Rozas J. 2002. Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolution*, 19: 2092-2100.
- Rand DM. 1996. Neutrality tests of molecular markers and the connection between DNA polymorphism, demography, and conservation biology. *Conservation Biology*, 10: 665-671.

- Reynard NS, Prudhomme C y Crooks SM. 2001. The flood characteristics of large UK Rivers: potential effects of changing climate and land use. *Climatic Change*, 48: 343-359.
- Rice WR. 1989. Analyzing tables of statistical tests. *Evolution*, 43: 223-235.
- Rödger D, Kielgast J, Bielby J, Schidtlein S, Bosch J, Garner TWJ, Veith M, Fisher MC y Lötters S. 2009. Global amphibian extinction risk assessment for the panzootic chytrid fungus. *Diversity*, 1: 52-66.
- Rogers AR y Harpending H. 1992. Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution*, 9: 552-569.
- Roman J. 2006. Diluting the founder effect: cryptic invasions expand a marine invader's range. *Proceedings of the Royal Society of London, Series B*, 278: 2453-2459.
- Sanjur OI, Carmona JA y Doadrio I. 2003. Evolutionary and biogeographical patterns within Iberian populations of the genus *Squalius* inferred from molecular data. *Molecular Phylogenetics and Evolution*, 29: 20-30.
- Slatkin M y Hudson RR. 1991. Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing comparisons. *Genetics*, 129: 555-562.
- Snoj A, Maric S, Berrebi P, Crivelli AJ, Shumka S y Susnik S. 2009. Genetic architecture of trout from Albania as revealed by mtDNA control region variation. *Genetics Selection Evolution*, 41(1): 22.
- Sousa-Santos C, Robalo JI, Francisco SM, Carrapato C, Cardoso AC y Doadrio I. 2014. Metapopulations in temporary streams – the role of drought-flood cycles in promoting high genetic diversity in a critically endangered freshwater fish and its consequences for the future. *Molecular Phylogenetics and Evolution*, 80: 281-296.
- Swartz ER, Skelton, PH y Bloomer P. 2007. Sea-level changes, river capture and the evolution of populations of the Eastern Cape and fiery redbins (*Pseudobarbus after* and *Pseudobarbus phlegethon*, Cyprinidae) across multiple river systems in South Africa. *Journal of Biogeography*, 34: 2086-2099.
- Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, 123: 585-595.
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M y Kumar S. 2011. MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary distance, and Maximum Parsimony methods. *Molecular Biology and Evolution*, 28: 2731-2739.
- Thompson JD, Higgins DG y Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, 22(22): 4673-4680.
- Tockner K, Bunn S, Gordon C, Naiman RJ, Quinn GP y Stanford JA. 2008. Flood plains: critically threatened ecosystems. En: Polunin NVC (ed.). *Aquatic Ecosystems. Trends and Global Prospects*. Cambridge University Press, Cambridge, UK. pp. 41-65.
- Underwood EC, Viers JH, Klausmeyer KR, Cox RL y Shaw MR. 2009. Threats and biodiversity in the Mediterranean biome. *Diversity and Distributions*, 15: 188-197.
- Unmack PJ, Bagley JC, Adams M, Hammer MP y Johnson JB. 2012. Molecular phylogeny and phylogeography of the Australian freshwater fish genus *Galaxiella*, with an emphasis on dwarf *Galaxias* (*G. pusilla*). *PloS One*, 7: e38433.
- Waelbrock C, Labeyrie L. Michel E, 2002. Sea-level and deep water temperature changes derived from benthic foraminifera isotopic records. *Quaternary Science Review*, 21: 295-305.
- Weir BS y Cockerham CC. 1984. Estimating F-statistics for the analysis of population structure. *Evolution*, 38: 1358-1370.

SUPPLEMENTARY MATERIAL

Table S1. *Squalius valentinus* sampling locations, number of specimens examined and Genbank accession numbers.

Population (Basin)	Locality	Utm coordinates	Number of individuals	Population label	GenBank Accession numbers (Haplotype or Allele Frequency)
Mijares (Mijares)	1 Mijares R, Los Cantos, Teruel	30T 703784 4443460	5 (<i>MT-CYTB</i>) 5 (<i>RAG1</i>)	MIJ	<i>MT-CYTB</i> : KR871722 (1), KR871738 (4) <i>RAG1</i> : KR871769 (8), KR871773 (2)
Mijares (Mijares)	2 Mijares R, Espadilla, Teruel	30T 725684 4434231	1 (<i>MT-CYTB</i>) 1 (<i>RAG1</i>)	MIJ	<i>MT-CYTB</i> : KR871722 (1) <i>RAG1</i> : KR871769 (1), KR871773 (1)
Mijares (Mijares)	3 Mijares R, Olba, Teruel	30T 701228 4445227	19 (<i>MT-CYTB</i>) 14 (<i>RAG1</i>)	MIJ	<i>MT-CYTB</i> : KR871722 (7), KR871738 (12) <i>RAG1</i> : KR871769 (18), KR871773 (9), KR871775 (1)
Albentosa (Mijares)	Albentosa R, Albentosa, Teruel	30 T 689749 4441736	26 (<i>MT-CYTB</i>) 24 (<i>RAG1</i>)	ALBE	<i>MT-CYTB</i> : KR871722 (5), KR871738 (21) <i>RAG1</i> : KR871769 (35), KR871773 (4), KR871779 (1), KR871780 (1), KR871781 (1), KR871782 (2), KR871783 (4)
Carbo (Mijares)	Carbo R, Tributary of Villahermosa R. Villahermosa del Río, Castellón	30T 719791 4453224	5 (<i>MT-CYTB</i>) 5 (<i>RAG1</i>)	CAR	<i>MT-CYTB</i> : KR871722 (1), KR871738 (3), KR871739 (1) <i>RAG1</i> : KR871769 (7), KR871773 (1), KR871775 (2)
Villahermosa (Mijares)	Villahermosa R, Cedramán, Castellón	30T 0722795 4449102	36 (<i>MT-CYTB</i>) 31 (<i>RAG1</i>)	VIL	<i>MT-CYTB</i> : KR871722 (33), KR871738 (3) <i>RAG1</i> : KR871769 (48), KR871773 (11), KR871794 (1), KR871795 (2)
Turia 1 (Turia)	Turia R, Villamarchante, Valencia	30S 704229 4382466	5 (<i>MT-CYTB</i>)	TUR	<i>MT-CYTB</i> : KR871721 (1), KR871722 (3),

			5 (RAG1)		KR871745 (1) RAG1: KR871769 (6), KR871773 (4)
Turia 2 (Turia)	Turia R, Gestalgar, Valencia	30S 686112 4385961	5 (MT-CYTB) 5 (RAG1)	TUR	MT-CYTB: KR871722 (1), KR871731 (1), KR871748 (3) RAG1: KR871769 (9), KR871773 (1)
Turia 3 (Turia)	Turia R, Calles, Valencia	30S 673683 4399244	5 (MT-CYTB) 5 (RAG1)	TUR	MT-CYTB: KR871738 (2), KR871745 (2), KR871746 (1) RAG1: KR871769 (5), KR871773 (3), KR871778 (2)
Turia 4 (Turia)	Turia R, Chulilla, Valencia	30S 680900 4391748	20 (MT-CYTB) 13 (RAG1)	TUR	MT-CYTB: KR871722 (13), KR871748 (7) RAG1: KR871769 (20), KR871773 (6)
Tuéjar 1 (Turia)	Tuéjar R (Source), Valencia,	30S 667961 4403291	5 (MT-CYTB) 5 (RAG1)	TUE	MT-CYTB: KR871745 (3), KR871746 (2) RAG1: KR871769 (10)
Tuéjar 2 (Turia)	Tuéjar R, Calles, Valencia	30 S 673625 4399753	23 (MT-CYTB) 22 (RAG1)	TUE	MT-CYTB: KR871722 (15), KR871738 (1), KR871748 (5), KR871758 (1), KR871759 (1), KR871760 (2) RAG1: KR871769 (32), KR871773 (12)
Júcar 1 (Júcar)	Júcar R (Rambla Pampanera), Cortés de Pallás, Valencia	30S 677574 4345586	5 (MT-CYTB) 7 (RAG1)	JUC	MT-CYTB: KR871722 (3), KR871729 (1), KR871730 (1) RAG1: KR871769 (13), KR871773 (1)
Júcar 2 (Júcar)	Júcar R, Millares, Valencia	30S 692227 4345386	5 (MT-CYTB) 4 (RAG1)	JUC	MT-CYTB: KR871722 (3), KR871732 (1), KR871733 (1) RAG1: KR871769 (8)
Júcar 3 (Júcar)	Júcar R (irrigation pond), Antela, Valencia,	30S 708209 4328302	5 (MT-CYTB) 5 (RAG1)	JUC	MT-CYTB: KR871740 (2), KR871741 (2), KR871742 (1) RAG1: KR871769 (10)
Albaida (Júcar)	Albaida R, Genovés, Valencia	30S 717852 4316704	4 (MT-CYTB)	ALBA	MT-CYTB: KR871722 (2), KR871748 (2)

				4 (RAG1)		RAG1: KR871769 (7), KR871773 (1)
Barranco del agua (Júcar)	Barranco del agua, Tributary of Júcar basin, Jarafuel, Valencia	30S 664334 4334018	16 (MT-CYTB) 7 (RAG1)	AGUA		MT-CYTB: KR871722 (16) RAG1: KR871769 (12), KR871785 (2)
Cabriel 1 (Júcar)	Cabriel R, Requena, Valencia	30S 663373 4372481	7 (MT-CYTB) 7 (RAG1)	CAB		MT-CYTB: KR871722 (7) RAG1: KR871769 (14)
Cabriel 2 (Júcar)	Cabriel R (Rambla Albosa), Venta del Moro, Valencia	30S 640897 4371913	2 (MT-CYTB)	CAB		MT-CYTB: KR871720 (1), KR871722 (1)
Cabriel 3 (Júcar)	Cabriel R Vadicañas, Cuenca.	30S 627085 4367169	3 (MT-CYTB)	CAB		MT-CYTB: KR871719 (2), KR871722 (1)
Grande (Júcar)	Grande R, Quesa, Valencia	30S 0695064 4333065	40 (MT-CYTB) 25 (RAG1)	GRA		MT-CYTB: KR871722 (16), KR871735 (3), KR871736 (1), KR871437 (1), KR871748 (2), KR871757 (17) RAG1: KR871769 (45), KR871773 (1), KR871775 (2), KR871793 (2)
Magro 1 (Júcar)	Magro R, Montroy, Valencia	30S 705579 4357255	7 (MT-CYTB) 7 (RAG1)	MAG		MT-CYTB: KR871722 (5), KR871734 (1), KR871735 (1) RAG1: KR871768 (2), KR871769 (10)
Magro 2 (Júcar)	Magro R, Turis, Valencia.	30S 697216 4362401	6 (MT-CYTB) 7 (RAG1)	MAG		MT-CYTB: KR871722 (5), KR871747 (1) RAG1: KR871769 (11), KR871773 (3)
Micena (Júcar)	Micena R, Otos, Valencia	30S 721722 4303673	32 (MT-CYTB) 27 (RAG1)	MIC		MT-CYTB: KR871748 (32) RAG1: KR871769 (51), KR871773 (1), KR871784 (2)
Sellent (Júcar)	Sellent R, Sellent, Valencia	30S 708812 4322980	5 (MT-CYTB) 5 (RAG1)	SEL		MT-CYTB: KR871722 (2), KR871734 (1), KR871735 (2) RAG1: KR871769 (6), KR871773 (4)
Verde (Júcar)	Verde R, Massalavés, Valencia	30S 714225 4335320	1 (MT-CYTB) 1 (RAG1)	VER		MT-CYTB: KR871722 (1) RAG1: KR871769 (2)
Albufera 1	Font del Forner spring, Sollana,	30S 725883 4350876	5 (MT-CYTB)	ALBU		MT-CYTB: KR871723 (4), KR871724 (1)

(Albufera)	Valencia		5 (RAG1)		RAG1: KR871769 (8), KR871771 (2)
Albufera (Albufera)	2 Font del Barret spring, Sollana, Valencia	30S 725883 4350876	33 (MT-CYTB) 18 (RAG1)	ALBU	MT-CYTB: KR871722 (14), KR871723 (1), KR871733 (1), KR871741 (5), KR871748 (6), KR871761 (1), KR871762 (1), KR871763 (1), KR871764 (1), KR871765 (1), KR871766 (1) RAG1: KR871769 (28), KR871770 (2), KR871773 (6)
Bullent (Serpis)	Bullent R, Pego, Alicante	30S 752960 4307746	11 (MT-CYTB) 8 (RAG1)	BUL	MT-CYTB: KR871722 (4), KR871725 (1), KR871726 (1), KR871727 (3), KR871728 (1), KR871767 (1) RAG1: KR871769 (14), KR871772 (2)
Serpis (Serpis)	Serpis R, Barranco de la Encantada. Beniarrés, Alicante	30S 0730211 4300889	31 (MT-CYTB) 27 (RAG1)	SER	MT-CYTB: KR871722 (9), KR871727 (19), KR871743 (1), KR871744 (1), KR871750 (1) RAG1: KR871769 (42), KR871772 (3), KR871776 (2), KR871777 (3), KR871791 (2), KR871792 (2)
Vinalopó (Vinalopó)	Vinalopó R, Banyeres de Mariola, Alicante	30S 703293 4287051	30 (MT-CYTB) 27 (RAG1)	VIN	MT-CYTB: KR871722 (29), KR871756 (1) RAG1: KR871769 (52), KR871790 (2)
Monnegre (Monnegre)	Monnegre R, Tibi, Alicante	30S 709792 4268351	37 (MT-CYTB) 35 (RAG1)	MON	MT-CYTB: KR871722 (36), KR871749 (1) RAG1: KR871769 (62), KR871774 (2), KR871786 (2), KR871787 (2), KR871788 (2)
Algar (Algar)	Algar R, Callosa D'en Sarriá, Alicante	30S 752598 4282925	34 (MT-CYTB) 21 (RAG1)	ALG	MT-CYTB: KR871722 (20), KR 871738 (2), KR871751 (3), KR871752 (1), KR871753 (6), KR871754 (1), R871755 (1) RAG1: KR871769 (39), KR871773 (1), KR871789 (2)

Table S2. *Squalius valentinus* localities used in MAXENT analysis.

Locality	Longitude	Latitude
1	0.24	40.21
2	-0.00	38.85
3	-0.04	40.03
4	-0.06	38.86
5	-0.07	38.59
6	-0.07	38.66
7	-0.07	38.68
8	-0.13	39.94
9	-0.17	38.95
10	-0.17	39.04
11	-0.18	38.68
12	-0.18	38.86
13	-0.26	39.58
14	-0.28	39.04
15	-0.28	39.22
16	-0.29	38.86
17	-0.36	40.03
18	-0.36	40.12
19	-0.36	40.21
20	-0.39	39.22
21	-0.39	39.31
22	-0.40	39.04
23	-0.41	38.86
24	-0.42	38.50
25	-0.47	40.22
26	-0.49	39.77
27	-0.49	39.86
28	-0.50	39.50
29	-0.51	39.14
30	-0.52	38.87
31	-0.52	39.05
32	-0.53	38.51
33	-0.53	38.69
34	-0.59	40.13

35	-0.60	39.86
36	-0.61	39.59
37	-0.62	39.32
38	-0.62	39.41
39	-0.62	39.50
40	-0.63	39.05
41	-0.63	39.14
42	-0.64	38.69
43	-0.64	38.78
44	-0.64	38.87
45	-0.72	39.86
46	-0.73	39.59
47	-0.74	39.14
48	-0.74	39.23
49	-0.74	39.32
50	-0.74	39.41
51	-0.75	38.87
52	-0.75	39.05
53	-0.82	39.32
54	-0.85	39.32
55	-0.85	39.41
56	-0.85	39.59
57	-0.86	39.23
58	-0.96	39.51
59	-0.96	39.60
60	-0.96	39.69
61	-0.97	39.24
62	-0.97	39.33
63	-0.97	39.42
64	-1.07	39.78
65	-1.08	39.51
66	-1.08	39.69
67	-1.09	39.06
68	-1.09	39.15
69	-1.09	39.33
70	-1.20	39.33
71	-1.20	39.42

72	-1.20	39.51
73	-1.31	39.60
74	-1.32	39.33
75	-1.43	39.33
76	-1.43	39.42
77	-1.55	39.42
78	-1.55	39.51
79	-0.66	38.71
80	-1.10	39.49
81	-1.14	39.29
82	-1.21	39.43
83	-1.36	39.49
84	-1.52	39.44
85	-0.61	39.34
86	-0.73	39.60
87	-1.21	39.43
88	-0.71	39.39
89	-0.94	39.24
90	-0.77	39.24
91	-0.59	39.08
92	-0.59	39.03
93	-0.74	39.12
94	-0.39	40.16
95	-0.38	40.16
96	-0.62	39.57
97	-0.83	39.60
98	-0.97	39.73
99	-0.89	39.66
100	-1.04	39.76
101	-0.97	39.73
102	-0.38	39.28
103	-0.08	38.88
104	-0.52	39.14
105	-0.35	38.83
106	-0.61	40.12
107	-0.64	40.13
108	-0.10	38.66

109	-0.44	38.85
110	-0.59	38.54
111	-0.49	38.97
112	-0.42	40.20
113	-1.10	39.14
114	-0.77	40.10

Table S3. Genetic diversity of populations in each river, major basins, and the complete dataset based on mitochondrial cytochrome *b* and nuclear *RAG1* genes.

Population	n	h	p (SD)	HD (SD)	S	k
Complete dataset	478	49	0.0011 (0.000007)	0.689 (0.022)	63	1.242
	762	28	0.00033 (0.00003)	0.301 (0.021)	29	0.342
Mijares basin	92	3	0.00054 (0.00008)	0.515 (0.014)	5	0.591
	160	10	0.00046 (0.00006)	0.436 (0.043)	9	0.11
Mijares R.	25	3	0.00044 (0.00005)	0.480 (0.060)	1	0.480
	40	3	0.00043 (0.00007)	0.465 (0.062)	2	0.481
Villahermosa R.	36	2	0.00022 (0.00007)	0.157 (0.077)	1	0.157
	62	4	0.00038 (0.00059)	0.374 (0.068)	4	0.413
Carbo R.	5	3	0.00182 (0.0000007)	0.700 (0.218)	5	2
	10	3	0.00051 (0.00019)	0.511 (0.164)	2	0.556
Albentosa R.	26	2	0.00029 (0.00009)	0.323 (0.097)	1	0.323
	48	7	0.00055 (0.00013)	0.461 (0.086)	6	0.593
Turia basin	65	10	0.00087 (0.00010)	0.700 (0.048)	10	0.952
	110	3	0.00037 (0.00004)	0.392 (0.044)	2	0.4
Turia R.	35	7	0.00084 (0.00013)	0.689 (0.060)	7	0.921
	56	3	0.00042 (0.00006)	0.434 (0.061)	2	0.452
Tuejar R.	30	8	0.00090 (0.00016)	0.724 (0.076)	7	0.993
	54	2	0.00032 (0.00006)	0.352 (0.064)	1	0.352
Albufera L.	38	12	0.00147 (0.00021)	0.821 (0.046)	14	1.616
	46	4	0.00051 (0.00014)	0.375 (0.084)	4	0.549
Júcar basin	131	13	0.00152 (0.00014)	0.693 (0.029)	16	1.664
	210	7	0.00018 (0.00004)	0.188 (0.036)	6	0.195
Júcar R.	10	5	0.00091 (0.00032)	0.667 (0.163)	5	1
	22	2	0.00008 (0.00007)	0.091 (0.091)	1	0.091
Júcar R. (Irrigation pond in Antela)	5	3	0.0046 (0.00036)	0.800 (0.164)	3	1.6
	-	-	-	-	-	-
Cabriel R.	12	3	0.00052 (0.00021)	0.439 (0.158)	2	0,356
	14	1	-	-	-	-
Magro R.	13	4	0.00098 (0.00050)	0.423 (0.164)	7	1.077
	26	3	0.00033 (0.00011)	0.342 (0.110)	2	0.110
Grande R.	40	6	0.00177 (0.00014)	0.667 (0.046)	7	1.945
	50	4	0.00018 (0.00007)	0.190 (0.0073)	3	0.197
Sellent R.	5	3	0.00291 (0.00073)	0.800 (0.164)	6	3.2
	10	2	0.00049 (0.00009)	0.533 (0.095)	1	0.533
Albaida R.	4	2	0.00061 (0.00019)	0.667 (0.204)	1	0.667
	8	2	0.00023 (0.00017)	0.250 (0.180)	1	0.25
Micena R.	32	1	-	-	-	-
	54	3	0.00010 (0.00005)	0.108 (0.057)	2	0.11
Barranco del Agua	16	1	-	-	-	-

	-	-	-	-	-	-
Bullent basin	11	6	0.00149 (0.00036)	0.836 (0.089)	7	1.636
	16	2	0.00021 (0.00012)	0.233 (0.126)	1	0.233
Serpis basin	31	5	0.00068 (0.00018)	0.555 (0.075)	6	0.748
	54	6	0.0004 (0.00009)	0.392 (0.083)	5	0.432
Vinalopó basin	30	2	0.00012 (0.00011)	0.067 (0.061)	2	0.133
	54	2	0.00007 (0.00004)	0.073 (0.048)	1	0.115
Monnegre basin	37	2	0.00005 (0.00005)	0.054 (0.050)	1	0.054
	70	5	0.00026 (0.00096)	0.215 (0.065)	4	0.282
Algar basin	36	7	0.00074 (0.00014)	0.660 (0.076)	6	0.813
	42	3	0.00013 (0.00007)	0.138 (0.071)	2	0.141

Table S4. Pairwise Φ_{ST} comparisons (above diagonal) among *Squalius valentinus* populations based on cytochrome *b* gene.

	Mijares	Turia	Albufera	Júcar	Serpis	Bullent	Algar	Monnegre	Vinalopo
Mijares	-								
Turia	0.27366	-							
Albufera	0.26543	0.06278	-						
Júcar	0.23739	0.06633	0.08777	-					
Serpis	0.38597	0.38597	0.32124	0.29709	-				
Bullent	0.20572	0.20572	0.14250	0.16790	0.05834	-			
Algar	0.19651	0.08703	0.09207	0.11289	0.40579	0.19843	-		
Monnegre	0.33539	0.08342	0.09251	0.09845	0.57426	0.36969	0.08230	-	
Vinalopó	0.31472	0.07326	0.07791	0.09174	0.52672	0.30365	0.06687	0.00283	-

Table S5. Demographic characteristics of each population of rivers and major basins and the complete dataset for mitochondrial cytochrome *b* gene.

Populations	R2 (<i>P</i>)	F _s (<i>P</i>)	D (<i>P</i>)	<i>r</i> (<i>P</i>)	Time since expansion (kya)
Entire dataset	0.010 (0.007)	-60.410 (0.0001)	-2.448 (<0.001)	-7.345 (>0.10)	37 200-70 600
Mijares	0.097 (0.521)	1.109 (0.764)	-0.861 (>0.10)	-2.836 (<0.05)	-
Turia	0.0511 (0.063)	-5.097 (0.003)	-1.498 (>0.10)	-2.052 (>0.05)	57 700-109 500
Albufera	0.0523 (0.016)	-7.759 (0.001)	-1.704 (>0.05)	-3.032 (>0.05)	101 000-191 900
Júcar	0.0510 (0.143)	-3.445 (0.089)	-1.172 (>0.10)	-2.587 (<0.05)	-
Bullent	0.121 (0.014)	-2.183 (0.032)	-1.230 (>0.10)	-1.723 (>0.10)	99 200-188 000
Serpis	0.102 (0.228)	-1.359 (0.141)	-1.420 (>0.10)	-2.692 (<0.05)	31 400-59 700
Vinalopó	0.179 (0.548)	-0.396 (0.371)	-1.507 (>0.10)	-2.381 (>0.10)	-
Monnegre	0.162 (0.523)	-1.385 (0.238)	-1.130 (>0.10)	-1.817 (>0.10)	-
Algar	0.087 (0.082)	-2.422 (0.466)	-1.033 (>0.10)	-0.776 (>0.10)	-

Table S6. Long-term (historical) migration analysis of *Squalius valentinus* populations based on cytochrome *b* gene.

MIGRATING POPULATION			RECEIVING POPULATION								
	n	θ	MIJ	TUR	JUC	ALBU	SER	BUL	ALG	MON	VIN
MIJ	92	0.0006	-	M=17491 m=0.30	M=4892 m=0.08	M=2775 m=0.04	M=13341 m=0.24	M=8908 m=0.16	M=15675 m=0.26	M=10408 m=0.18	M=11035 m=0.18
TUR	65	0.0010	M=11842 m=0.21	-	M=10842 m=0.19	M=5108 m=0.09	M=13208 m=0.23	M=12142 m=0.21	M=10142 m=0.18	M=11825 m=0.20	M=11141 m=0.19
JUC	131	0.0010	M=9525 m=0.16	M=12408 m=0.21	-	M=6092 m=0.11	M=11408 m=0.20	M=9175 m=0.16	M=6441 m=0.11	M=6525 m=0.11	M=7869 m=0.14
ALBU	38	0.0625	M=12958 m=0.22	M=17158 m=0.30	M=16742 m=0.29	-	M=15091 m=0.26	M=18025 m=0.31	M=16575 m=0.29	M=17375 m=0.30	M=15479 m=0.27
SER	31	0.0006	M=9225 m=0.16	M=5175 m=0.09	M=3358 m=0.06	M=2608 m=0.04	-	M=17941 m=0.31	M=5441 m=0.09	M=5342 m=0.09	M=5892 m=0.10
BUL	11	0.056	M=13108 m=0.23	M=15241 m=0.26	M=12642 m=0.22	M=10058 m=0.17	M=15341 m=0.27	-	M=15258 m=0.26	M=15975 m=0.28	M=15582 m=0.27
ALG	36	0.0567	M=13075 m=0.23	M=15875 m=0.28	M=13475 m=0.23	M=12241 m=0.21	M=14425 m=0.25	M=16925 m=0.29	-	M=17608 m=0.31	M=15708 m=0.27
MON	37	0.0067	M=13208 m=0.23	M=16841 m=0.29	M=14375 m=0.25	M=13325 m=0.23	M=15558 m=0.27	M=16608 m=0.29	M=16208 m=0.28	-	M=17241 m=0.30
VIN	30	0.0571	M=13591 m=0.24	M=16541 m=0.29	M=14358 m=0.25	M=11625 m=0.20	M=15591 m=0.27	M=17941 m=0.31	M=16375 m=0.28	M=16792 m=0.29	-

Table S7. Sum of Bayesian posterior probabilities of regression models that include a given factor.

VARIABLE	LIG	LGM	CURRENT PERIOD	FUTURE (2080)
TEMPERATURE (ALL VARIABLES)				
BIO1: Annual mean temperature	0.322	0.473	0.484	0.480
BIO2: Mean diurnal range	0.151	0.471	0.502	0.428
BIO3: Isothermality	0.390	0.459	0.511	0.454
BIO4: Temperature seasonality	0.311	0.573	0.494	0.477
BIO5: Maximum temperature of the warmest month	0.423	0.471	0.466	0.489
BIO6: Minimum temperature of the coldest month	0.356	0.481	0.477	0.466
BIO7: Temperature annual range;	0.363	0.702	0.472	0.443
BIO8: Mean temperature of the wettest quarter	0.192	0.624	0.462	0.467
BIO9: Mean temperature of the driest quarter;	0.331	0.485	0.473	0.483
BIO10: Mean temperature of the warmest quarter	0.324	0.482	0.510	0.462
BIO11: Mean temperature of the coldest quarter	0.349	0.492	0.478	0.470
PRECIPITATION (ALL VARIABLES)				
BIO12: Precipitation of the warmest month	0.422	0.310	0.462	0.446
BIO13: Precipitation of the wettest moth	0.442	0.292	0.610	0.494
BIO14: Precipitation of the driest month	0.492	0.345	0.458	0.501
BIO15: Precipitation seasonality	0.461	0.412	0.541	0.442
BIO16: Precipitation of the wettest quarter	0.447	0.340	0.465	0.439
BIO17: Precipitation of the driest quarter	0.473	0.389	0.465	0.457
BIO18: Precipitation of the warmest quarter	0.475	0.319	0.475	0.491
BIO19: Precipitation of the coldest quarter	0.460	0.310	0.461	0.449
SEASONALITY				
BIO4: Temperature seasonality	0.230	0.466	0.420	0.464
BIO7: Temperature annual range	0.581	0.266	0.265	0.590
BIO15: Precipitation seasonality	0.404	0.235	0.251	0.467
“EXTREME” CLIMATIC CONDITIONS				
BIO5: Maximum temperature of the warmest month	0.177	0.215	0.285	0.180
BIO6: Minimum temperature of the coldest month	0.864	0.276	0.274	0.669
BIO13: Precipitation of the wettest moth	0.124	0.214	0.251	0.114
BIO14: Precipitation of the driest month	0.243	0.202	0.197	0.400

Table S8. Sum of Bayesian posterior probabilities of regression models that include a given factor when only the five variables with the highest Bayesian probability values of temperature and precipitation are analysed.

VARIABLE	LIG	LGM	CURRENT PERIOD	FUTURE (2080)
TEMPERATURE (5 VARIABLES)				
BIO1: Annual mean temperature	-	-	0.790	0.474
BIO2: Mean diurnal range	-	-	0.510	-
BIO3: Isothermality	0.499	-	0.560	-
BIO4: Temperature seasonality	-	0.452	0.428	-
BIO5: Maximum temperature of the warmest month	0.465	-	-	0.450
BIO6: Minimum temperature of the coldest month	0.694	-	-	0.458
BIO7: Temperature annual range	0.616	0.454	-	-
BIO8: Mean temperature of the wettest quarter	-	0.429	-	-
BIO9: Mean temperature of the driest quarter;	-	0.569	-	0.421
BIO10: Mean temperature of the warmest quarter	-	0.471	0.797	-
BIO11: Mean temperature of the coldest quarter	0.570	0.426	-	0.529
PRECIPITATION (5 VARIABLES)				
BIO12: Precipitation of the warmest month	-	-	-	-
BIO13: Precipitation of the wettest moth	-	-	0.542	0.544
BIO14: Precipitation of the driest month	0.467	0.468	-	0.764
BIO15: Precipitation seasonality	0.393	0.425	0.542	-
BIO16: Precipitation of the wettest quarter	-	0.431	0.538	-
BIO17: Precipitation of the driest quarter	0.442	0.545	0.459	0.636
BIO18: Precipitation of the warmest quarter	0.449	0.802	0.463	0.430
BIO19: Precipitation of the coldest quarter	0.149	-	-	0.435

Appendix S1. Environmental variables analysed in GESTE software (online version).

Appendix S2. Genetic data input file used in GESTE software (online version).

Figure S1. Historical demography of *Squalius valentinus* based on Mismatch distributions (A) and Extended Bayesian Skyline plots (EBSP) of effective population size (N_e) over time (B).

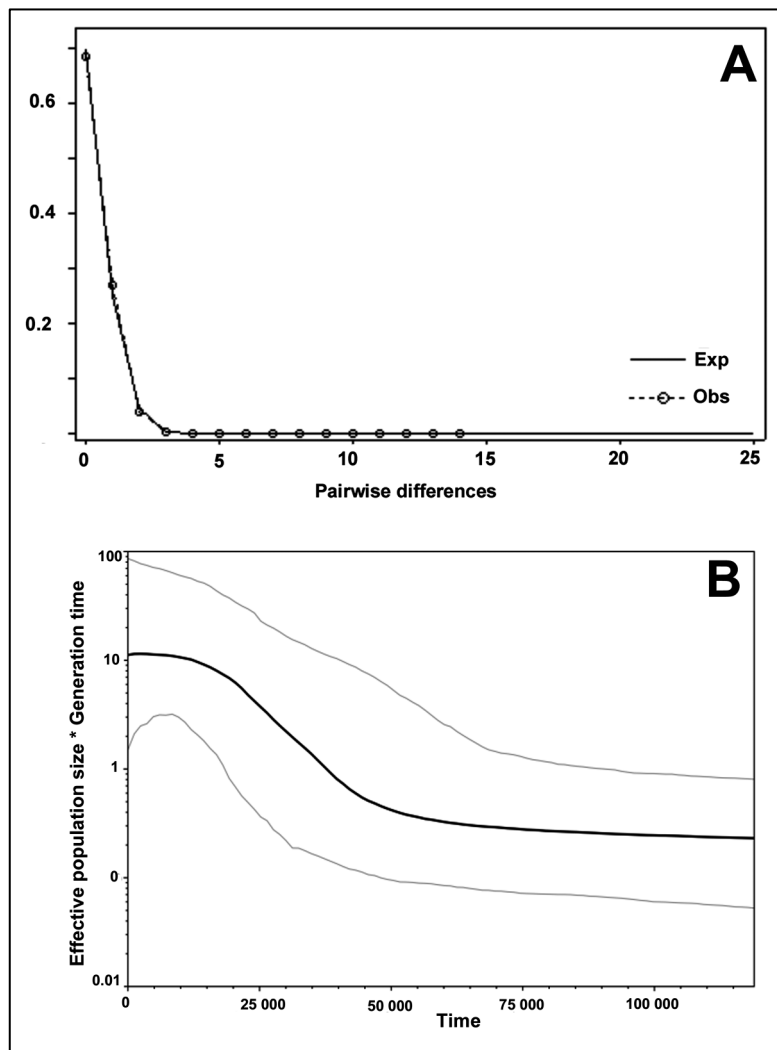


Figure S2. Mismatch distributions observed (dotted lines) and expected under a expansion model (continous line) for individual *Squalius valentinus* populations (A) and for the Júcar Basin (B).

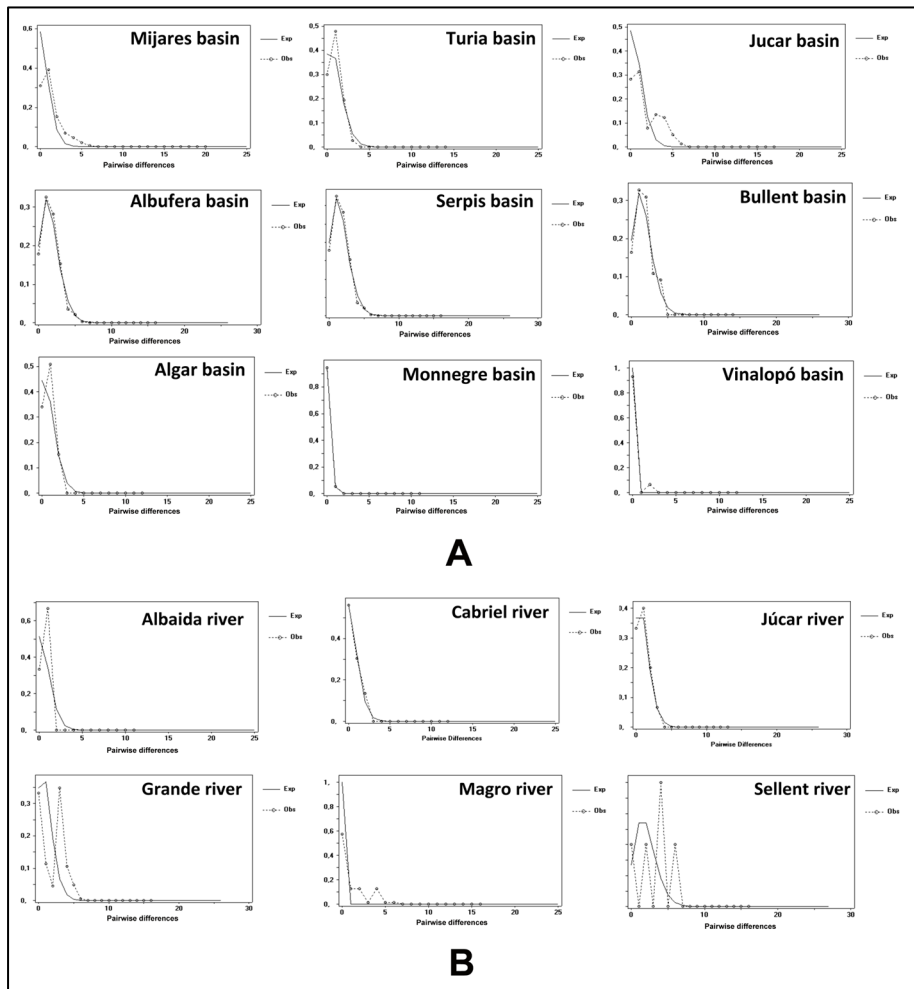
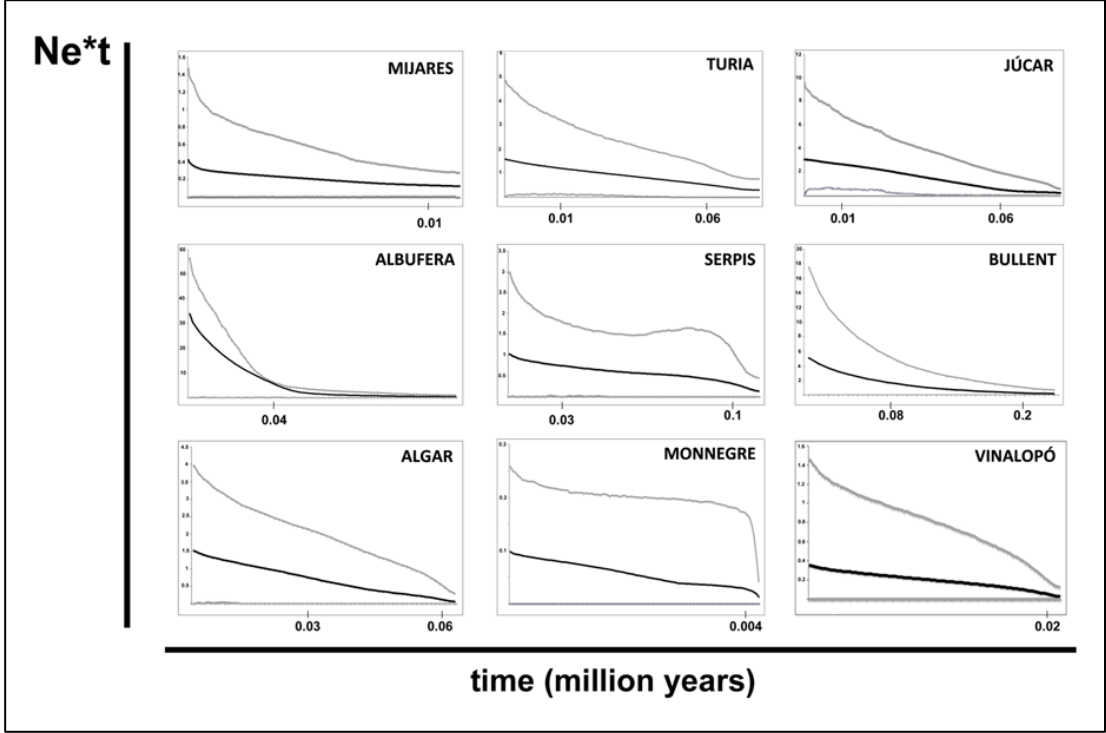


Figure S3. Historical demography of individual *Squalius valentinus* populations based on Extended Bayesian Skyline plot (EBSP) of effective population size (N_e) over time.



**Ancient mitochondrial capture as factor promoting
mitonuclear discordance in freshwater fishes: a case study
in the genus *Squalius* (Actinopterygii, Cyprinidae) in Greece**

Perea S, Vukic J, Sanda R y Doadrio I. Ancient mitochondrial capture as factor promoting mitonuclear discordance in freshwater fishes: a case study in the genus *Squalius* (Actinopterygii, Cyprinidae) in Greece. *PLoS One*, 11(12): e0166292.

DOI: 10.1371/journal.pone.0166292.

RESUMEN

La hibridación y la separación incompleta de linajes son dos factores que comúnmente confunden la filogenia y los procesos de especiación resultando en discordancias mitonucleares. La introgresión mitocondrial, un caso particular de hibridación, puede derivar en casos extremos en el reemplazamiento completo del genoma mitocondrial de una especie por otra (captura mitocondrial). En este estudio investigamos la introgresión mitocondrial ocurrida entre dos especies del género de ciprínido *Squalius* en la región occidental del Peloponeso griego utilizando datos morfológicos y moleculares. Como resultado encontramos evidencia de una introgresión mitocondrial completa de *S. keadicus* en dos poblaciones reconocidas como pertenecientes a *S. peloponensis* en las cuencas de los ríos Miras y Pamissos, y una divergencia del genoma mitocondrial de *S. keadicus* de la cuenca del río Evrotas respecto al genoma mitocondrial hallado en las poblaciones introgresadas datada en el Pleistoceno. Contactos secundarios entre cuencas es un posible factor responsable de la conexión entre especies y del evento de introgresión. Los análisis morfológicos apoyan la hipótesis de introgresión mitocondrial, ya que *S. keadicus* es morfológicamente diferente a las otras tres poblaciones adscritas a *S. peloponensis*, aunque diferencias significativas específicas fueron halladas entre las cuatro poblaciones. El aislamiento por barreras biogeográficas emergidas durante el Pleistoceno en la región occidental del Peloponeso es la probable fuente de evolución de los dos linajes mitocondriales recíprocamente monofiléticos hallados en el clado mitocondrial al que pertenece *S. keadicus*, y posiblemente también de las diferencias morfológicas halladas entre las cuatro poblaciones. Esto, junto con la falta de estructura en el genoma nuclear de las tres poblaciones reconocidas como *S. peloponensis*, sugiere un proceso de especiación incipiente en estas especies de *Squalius* de la región occidental del Peloponeso.

ABSTRACT

Hybridization and incomplete lineage sorting are common confounding factors in phylogeny and speciation resulting in mitonuclear disparity. Mitochondrial introgression, a particular case of hybridization may, in extreme cases, lead to replacement of the mitochondrial genome of one species with that of another (mitochondrial capture). We investigated mitochondrial introgression involving two species of the cyprinid genus *Squalius* in the western Peloponnese region of Greece using molecular and morphological data. We found evidence of complete mitochondrial introgression of *Squalius keadicus* into two populations recognized as *Squalius peloponensis* from the Miras and Pamissos River basins and a divergence of mitochondrial genomes of *S. keadicus* from the Evrotas basin from that of the introgressed populations dating from the Pleistocene. Secondary contact among basins is a possible factor in connection of the species and the introgression event. Morphological analyses support the hypothesis of mitochondrial introgression, as *S. keadicus* was different from the other three populations recognized as *S. peloponensis*, although significant differences were found among the four populations. Isolation by geographical barriers arose during Pleistocene in the western Peloponnese were the source of the evolution of the two reciprocally monophyletic subclades found in the *S. keadicus* mitochondrial clade, and the morphological differences found among the four populations. Along with the lack of structure in the nuclear genome in the three populations ascribed to *S. peloponensis*, this suggests an incipient speciation process occurring in these *Squalius* species in the western Peloponnese.

INTRODUCTION

The use of mitochondrial DNA for inferring phylogenetic or phylogeographic relationships has been traditionally preferred over the use of the nuclear genome, as mtDNA accurately reflects recent divergence patterns than do nuclear markers (Zink and Barraclough, 2008). Nevertheless, concordant patterns between mtDNA and nuclear DNA are not always observed in nature (Funk and Omland, 2003; Chan and Levin, 2005; Toews and Brelsford, 2012). Evolutionary processes such as hybridization and incomplete lineage sorting are common confounding factors in phylogeny and speciation and primary causes of mitonuclear discordance (Chan and Levin, 2005; Joly *et al.*, 2009; Toews and Brelsford, 2012). Incomplete lineage sorting occurs when the coalescence time of genes and speciation differ, i.e. when lineages fail to sort out at the same time at speciation happens (Maddison, 1997). Therefore, gene trees do not always represent the true relationships among taxa because different genes may have not branched at the same time (Degnan and Rosenberg, 2009; Edwards *et al.*, 2009). Distinguishing incomplete lineage sorting from hybridization can be difficult as these two processes may originate the same phylogenetic tree (Meng and Kubatko, 2009; McKey and Zink, 2010; Yu *et al.*, 2011). Nonetheless, among others, any biogeographic pattern is expected when incomplete lineage sorting is the confounding factor (Toews and Brelsford, 2012).

Hybridization between species involves mating between unrelated organisms regardless the taxonomic status and, in some cases, may lead to gene transfer, a common process in plants (Jan and Fernández-Martínez, 2002) or to very complex evolutionary processes occurring in freshwater fishes (Cunha *et al.*, 2004; Sousa-Santos *et al.*, 2006). Introgression is a particular case of hybridization and occurs when gene flow exists between populations that hybridize and hybrids backcross to one or both parental populations (Rhymer and Simberloff, 1996). Within this framework, mitochondrial capture is defined as complete mitochondrial introgression, a situation in which the mitochondrial genome of one species is replaced with that of another in an entire population, as a consequence of selective backcrossing of hybrids with one of the parent species (Funk and Omland, 2003; Chan and Levin, 2005; Weisrock *et al.*, 2005; Linnen and Farrel, 2007; Bryson *et al.*, 2014). In hybridization, the complete replacement of the mitochondrial DNA of one species with that of another is more common than is replacement of a portion of the nuclear genome, due to maternal inheritance and the four-fold smaller effective population size of mitochondrial DNA compared to nuclear DNA, implying that mtDNA will complete the process of lineage sorting more rapidly than will nuclear DNA (Ballard and Whitlock, 2004; Chan and Levin, 2005; Currat *et al.*, 2008; Zink and Barraclough, 2008).

Understanding evolutionary processes such as hybridization, mitochondrial introgression, and incomplete lineage sorting is of vital importance to inference reliable phylogenies and reconstruction of evolutionary history and speciation processes, especially in currently allopatric

species. Revealing mitochondrial introgression may be the only way to identify past hybridization of allopatric species (Weisrock *et al.*, 2005).

In general, mitonuclear discordance due to mitochondrial introgression occurs more frequently in certain taxonomic groups, including mammals and fish; whereas, in groups such as birds and butterflies, in which the female is heterogametic, mitochondrial introgression is usually reduced, conforming to Haldane's rule (Toews and Brelsford, 2012). Examples of the complete replacement of mitochondrial DNA are found in various taxa (Linnen and Farrel, 2007; Good *et al.*, 2008; Rabosky *et al.*, 2009; Marcková *et al.*, 2013; Bryson *et al.*, 2014; Willis *et al.*, 2014), including freshwater fishes (Freyhof *et al.*, 2005; Nevado *et al.*, 2009; Tang *et al.*, 2012). Within freshwater fishes the preponderance of hybridization is greater in cyprinids than is in other freshwater fish families such as salmonids or cichlids and several examples are observed (Costedoat *et al.*, 2007; Aboim *et al.*, 2010; Sinama *et al.*, 2013; Geiger *et al.*, 2014). A clear example of mitonuclear discordance has been described for the genus *Squalius* in the Greek Peloponnese. Previous studies focusing on a more general framework of phylogenetic relationships of *Squalius* hypothesised the presence of mitochondrial introgression to explain discrepancy found in the mitochondrial (DNA sequences) and nuclear (allozyme) phylogenetic position of *S. keadicus* (Stephanidis, 1971) and *S. peloponensis* (Valenncienes, 1844) (Doadrio and Carmona, 1998; Durand *et al.*, 2000). These two species, currently showing allopatric distribution, belong to two different evolutionary lineages recognized within the genus *Squalius*; *S. peloponensis* belongs to the Euroasiatic group while *S. keadicus* is included together with other small-sized species within the Mediterranean group. These two evolutionary lineages have been isolated since the Middle Miocene (Sanjur *et al.*, 2003; Perea *et al.*, 2010). *Squalius keadicus* is a small- to medium-sized fish endemic to the Evrotas Basin in the southwestern Peloponnese. *Squalius peloponensis* is a medium to large fish and has a broader distribution range in the western Peloponnese, encompassing several hydrological basins (Fig. 1) (Kottelat and Freyhof, 2007). The morphological distinction between both species is clear and several diagnostic morphometric and meristic characters are identified to distinguishing them from other Balkan *Squalius* species. Head length is 24-27% standard length in *S. peloponensis*. The posterior margin of the anal fin is different in both species, being straight in *S. keadicus* and slightly convex in *S. peloponensis*. Meristic diagnostic characters of *S. keadicus* show 44-49 scales on lateral line whereas in *S. peloponensis* this range is 40-44. The colouration of both species also differs; *S. keadicus* shows in live a conspicuous blackish stripe on flank from eye to caudal fin base and very dark scales in upper half of flank, which are not present in *S. peloponensis* (Kottelat and Freyhof, 2007).

With regards to the paleogeographical context in which the two species have evolved, the southwestern Peloponnese lies near the Hellenic Trench in the Eastern Mediterranean region, one of the most active areas in the Euroasiatic-African convergence zone owing to the

northward subduction of the African plate beneath the Aegean plate (Fountoulis *et al.*, 2014). The region has a complex tectonic history and numerous currently active northwestern-southeastern tectonic faults (Kokkalas *et al.*, 2006; Papanikolaou *et al.*, 2007; Fountoulis and Mariolakis, 2008; Fountoulis *et al.*, 2014). The formation of the hydrological network of the Greek Peloponnese was complex, characterized by Late Miocene-Quaternary compressional structures as well as Pliocene lacustrine-marine and Pleistocene terrestrial-fluvial sedimentary successions leading to the formation of fluvial terraces (Koukouvelas *et al.*, 2009; Atthanasas and Fountoulis, 2013; Fountoulis *et al.*, 2014). Such paleogeographic complexity may have promoted ancient contact among freshwater fauna inhabiting the basins of the western Peloponnese, resulting in hybridization events.

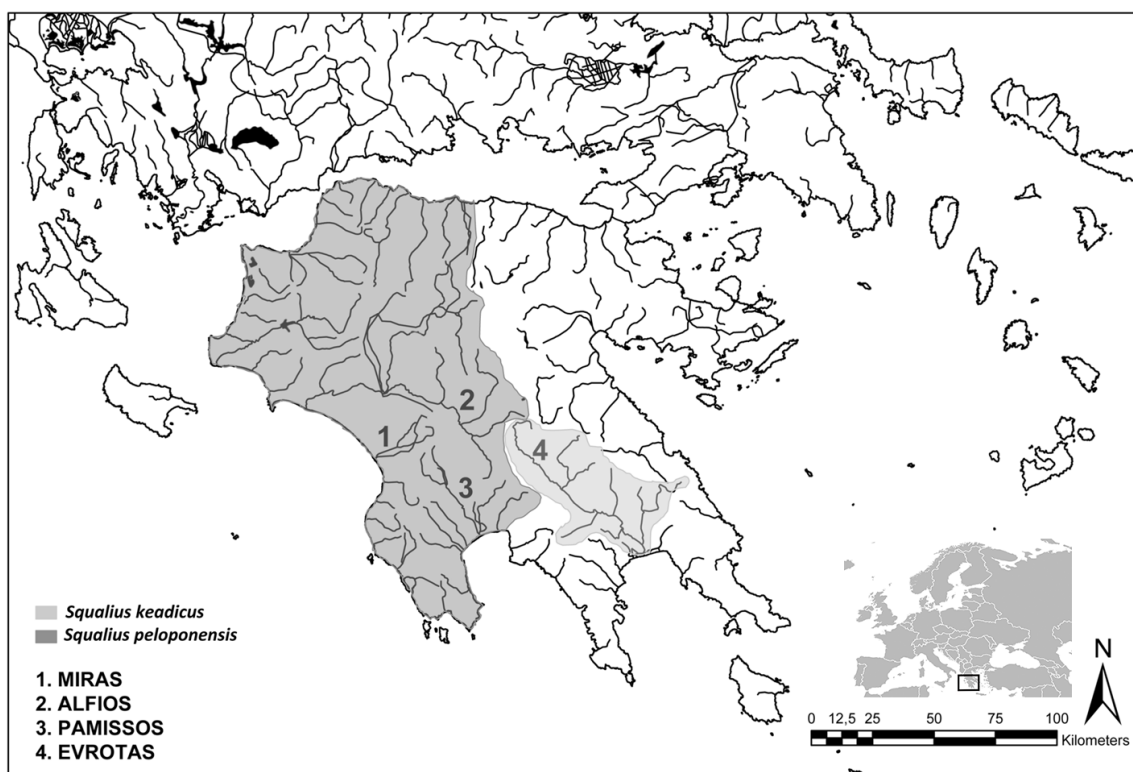


Fig. 1 Locality samples and distribution range of *S. keadicus* and *S. peloponensis* according to Kottelat & Freyhof (2007).

Within the framework of mitonuclear disparity, the aim of this study was to investigate the mitochondrial capture between two species of *Squalius* from the Greek Peloponnese using molecular and morphological analyses of three populations of *S. peloponensis* and one of *S. keadicus*. We also establish a temporal and biogeographical framework for the introgression episode.

MATERIAL AND METHODS

1. Sampling

Specimens of four populations of *Squalius* were collected by electrofishing from basins in the western Peloponnese region of Greece following the European regulations (*EN ISO 14011:200. Water quality - Sampling of fish with electrophising*) (Table 1; Fig. 1). Sampling was performed with authorised permission from the Greek Ministry of Environment, Energy and Climatic Change and after consulting the Animal Care and Use Committee of the National Museum of Prague (č.j.NM/2014/258). Immediately after capture, all individuals were anesthetized using Tricaine methanesulfonate (MS-222) to alleviate suffering. A piece of ventral fin sample (~3 mm² fin clips) was obtained from 61 fish belonging to the four Greek populations analysed (nine to *S. keadicus* and 52 to *S. peloponensis*). All tissues were preserved in 95% ethanol and stored at 4°C until its processing at laboratory for molecular analyses. These 61 specimens analysed for molecular markers were sacrificed for morphological analyses following humanely euthanization with overdose of MS-222 according to European Commission regulations (Directive 2010/63/EU). Death was confirmed after no gill movement was observed for at least 10 minutes. These new sampled specimens used for morphological purposes were preserved in a solution of 5 % formalin and after a week they were transferred to a 70% ethanol solution. These new-collected specimens have been deposited in the Fish Collection of the National Museum of Prague (Czech Republic). The remaining specimens used for morphological analyses in this study came from historical deposits of the Fish Collections of the Museo Nacional de Ciencias Naturales, CSIC, in Madrid (Spain). Voucher tissues of new collected specimens are deposited in the DNA Collections of the Museo Nacional de Ciencias Naturales in Madrid, CSIC, (Spain) and in the National Museum of Prague (Czech Republic).

The study included the Evrotas River population, this being the type locality of *S. keadicus* and currently the only known locality for the species, and three populations recognized as *S. peloponensis* from the Alfios, Miras, and Pamissos basins (Fig. 1). The type locality of *S. peloponensis* is uncertain, and only the Peloponnese peninsula (= Morea) is cited in its description (Cuvier and Valenciennes, 1844). Additional species of *Squalius* along its distribution range were included in the molecular analyses to assess the phylogenetic position of the four investigated populations (S1 Table). The genus *Petroleuciscus* was used as the outgroup based on previous phylogenetic studies (Sanjur *et al.*, 2003; Perea *et al.*, 2010).

2. DNA extraction, amplification and sequencing

Eight to 17 specimens per population were analysed (Table 1). Total genomic DNA was isolated using the commercial kit Biosprint 15 for blood and tissue (Qiagen). For each specimen the

complete mitochondrial cytochrome *b* (*MT-CYB*; 1140bp) gene and two nuclear genes, the first intron of the ribosomal *S7* gene (*S7*; final alignment including gaps = 977bp) and the third exon of the recombination activating gene (*RAG-1*; 1473bp), were amplified. The PCR protocols and primers followed (Perea *et al.*, 2010). The PCR products were purified by Exo-SAP-IT (USB, Cleveland, OH, USA) and directly sequenced by MacroGen Europe (Amsterdam, The Netherlands; <http://www.macrogen.com>) using a 3730XL DNA sequencer. All new sequences of haplotypes and alleles obtained in this study were deposited in the GenBank database (Accession Numbers: *MT-CYB*: KY070368-KY070424; *RAG1*: KY070425-KY070574; *S7*: KY070503-KY070574).

3. Phylogenetic analyses

For the nuclear markers, phylogenetic inference of independent alleles was also conducted on gamete phases using the PHASE algorithm (Stephens *et al.*, 2001; Stephens y Donelli, 2003) implemented in DnaSP v. 5.0. (Librado y Rozas, 2009) with a probability threshold of 0.9 to resolve alleles. Sequences were aligned using the default pairwise and multiple alignment parameters in Clustal W (Thompson *et al.*, 1994) implemented in MEGA v.7 (Kumar *et al.*, 2015). Alignments were later revised. Recombination of nuclear genes was assessed by the PHI test (Bruen *et al.*, 2006) implemented in SplitsTree v. 4.13 (Huson y Bryant, 2006). No traces of recombination were found in either *RAG-1* or *S7* genes ($p = 1.0$ for both genes).

For phylogenetic analyses, the best-fit model of evolution for each independent gene was estimated using jModelTest v2 (Darriba *et al.*, 2012) (S2 Table). Model parameters were used for subsequent phylogenetic analyses. Phylogenies based on Bayesian inference (BI) and maximum likelihood (ML) were constructed for independent genes in order to assess the phylogenetic position of each population based on mitochondrial and nuclear markers. BI was performed in MrBayes 3.2 (Ronquist *et al.*, 2012). Two simultaneous analyses were run for 10 million generations, each with four MCMC chains sampling every 1000 generations. Convergence was assessed using Tracer v.1.6 (Rambaut y Drummond, 2007). The 50% majority rule consensus tree was constructed after discarding the first 10% of generations as burn-in. For ML analyses, we used RaxML software implemented in the Trex-online server (Stamakis, 2006), employing the substitution model GTR+G+I for *MT-CYB* and GTR+G for nuclear genes and the rapid bootstrapping algorithm to estimate node confidence using different random seeds (1000 replicates) (Stamakis *et al.*, 2007).

Uncorrected-*p* genetic distances between mitochondrial haplotypes and between nuclear phased alleles were estimated using the MEGA v.7 software (Kumar *et al.*, 2015).

Table 1. Specimens used in molecular and morphological analyses and their voucher numbers. GenBank accession numbers for all analysed genes and labels in phylogenetic trees. All these specimens have been used in morphological analyses.

Species/Population (Voucher number)	Geographic coordinates	MT-CYB	RAG1	S7	Label in phylogenetic trees
<i>Squalius keadicus</i> /Evrotas (AJ252820)*	-	AJ252820	-	-	<i>Squalius keadicus</i> 1
<i>Squalius keadicus</i> /Evrotas (AF090760)*	36° 51' 15.6" N, 22° 40' 29.8" E	AF090760	-	-	<i>Squalius keadicus</i> 2
<i>Squalius keadicus</i> /Evrotas (MNCN_ICTIO 123.875)	36° 51' 15.6" N, 22° 40' 29.8" E	HM560185	HM560583	HM560572	<i>Squalius keadicus</i> 3
<i>Squalius keadicus</i> /Evrotas (NMP PV6 F1556)	37° 5' 28.71" N, 22° 25' 38.74" E	KY070419	KY070433 (a); KY070434 (b)	KY070547 (a); KY070548 (b)	<i>Squalius keadicus</i> 4
<i>Squalius keadicus</i> /Evrotas (NMP PV6 F1557)	37° 5' 28.71" N, 22° 25' 38.74" E	KY070420	KY070435 (a); KY070436 (b)	KY070549 (a); KY070550 (b)	<i>Squalius keadicus</i> 5
<i>Squalius keadicus</i> /Evrotas (NMP PV6 F1566)	37° 5' 28.71" N, 22° 25' 38.74" E	KY070421	KY070425 (a); KY070426 (b)	KY070503 (a); KY070504 (b)	<i>Squalius keadicus</i> 6
<i>Squalius keadicus</i> /Evrotas (NMP PV6 F1567)	37° 5' 28.71" N, 22° 25' 38.74" E	KY070422	KY070427 (a); KY070428 (b)	KY070505 (a); KY070506 (b)	<i>Squalius keadicus</i> 7
<i>Squalius keadicus</i> /Evrotas (NMP PV6 F1570)	37° 5' 28.71" N, 22° 25' 38.74" E	KY070423	KY070429 (a); KY070430 (b)	KY070551 (a); KY070552 (b)	<i>Squalius keadicus</i> 8
<i>Squalius keadicus</i> /Evrotas (NMP PV6 F1571)	37° 5' 28.71" N, 22° 25' 38.74" E	KY070424	KY070431 (a); KY070432 (b)	KY070553 (a); KY070554 (b)	<i>Squalius keadicus</i> 9
<i>Squalius peloponensis</i> /Alfios (MNCN_ICTIO 120.548)	37° 28' 21.7" N, 22° 04' 33.9" E	KY070368	-	KY070515 (a); KY070516 (b)	Alfios 1
<i>Squalius peloponensis</i> /Alfios (MNCN_ICTIO 120.559)	37° 28' 21.7" N, 22° 04' 33.9" E	KY070372	KY070497 (a); KY070498 (b)	-	Alfios 2
<i>Squalius peloponensis</i> /Alfios (MNCN_ICTIO 120.563)	37° 28' 21.7" N, 22° 04' 33.9" E	KY070374	KY070501 (a); KY070502 (b)	KY070523 (a); KY070524 (b)	Alfios 3

<i>Squalius peloponensis</i> /Alfios (MNCN_ICTIO 120.553)	37° 28' 21.7" N, 22° 04' 33.9" E	KY070369	KY070491 (a); KY070492 (b)	KY070517 (a); KY070518 (b)	Alfios 4
<i>Squalius peloponensis</i> /Alfios (MNCN_ICTIO 120.555)	37° 28' 21.7" N, 22° 04' 33.9" E	KY070370	-	-	Alfios 5
<i>Squalius peloponensis</i> /Alfios (MNCN_ICTIO 120.556)	37° 28' 21.7" N, 22° 04' 33.9" E	KY070371	KY070493 (a); KY070494 (b)	KY070519 (a); KY070520 (b)	Alfios 6
<i>Squalius peloponensis</i> /Alfios (MNCN_ICTIO 120.560)	37° 28' 21.7" N, 22° 04' 33.9" E	KY07037	KY070499 (a); KY070500 (b)	KY070521 (a); KY070522 (b)	Alfios 7
<i>Squalius peloponensis</i> /Alfios (MNCN_ICTIO 120.567)	37° 28' 21.7" N, 22° 04' 33.9" E	-	KY070495 (a); KY070496 (b)	-	Alfios 8
<i>Squalius peloponensis</i> /Miras (NMP PV6 F1470)	37° 17' 34.2" N, 21° 42' 7.86" E	KY070407	KY070467 (a); KY070468 (b)	KY070525 (a); KY070526 (b)	Miras 1
<i>Squalius peloponensis</i> /Miras (NMP PV6 F1471)	37° 17' 34.2" N, 21° 42' 7.86" E	KY070408	KY070469 (a); KY070470 (b)	KY070527 (a); KY070528 (b)	Miras 2
<i>Squalius peloponensis</i> /Miras (NMP PV6 F1472)	37° 17' 34.2" N, 21° 42' 7.86" E	KY070409	KY070471 (a); KY070472 (b)	-	Miras 3
<i>Squalius peloponensis</i> /Miras (NMP PV6 G210)	37° 17' 34.2" N, 21° 42' 7.86" E	KY070410	KY070473 (a); KY070474 (b)	KY070555 (a); KY070556 (b)	Miras 4
<i>Squalius peloponensis</i> /Miras (NMP PV6 G211)	37° 17' 34.2" N, 21° 42' 7.86" E	KY070411	KY070475 (a); KY070476 (b)	KY070557 (a); KY070558 (b)	Miras 5
<i>Squalius peloponensis</i> /Miras (NMP PV6 G212)	37° 17' 34.2" N, 21° 42' 7.86" E	KY070412	KY070477 (a); KY070478 (b)	KY070559 (a); KY070560 (b)	Miras 6
<i>Squalius peloponensis</i> /Miras (NMP PV6 G213)	37° 17' 34.2" N, 21° 42' 7.86" E	KY070413	KY070479 (a); KY070480 (b)	KY070561 (a); KY070562 (b)	Miras 7
<i>Squalius peloponensis</i> /Miras (NMP PV6 G214)	37° 17' 34.2" N, 21° 42' 7.86" E	KY070414	KY070481 (a); KY070482 (b)	KY070563 (a); KY070564 (b)	Miras 8
<i>Squalius peloponensis</i> /Miras (NMP PV6 G215)	37° 17' 34.2" N, 21° 42' 7.86" E	KY070415	KY070483 (a); KY070484 (b)	KY070565 (a); KY070566 (b)	Miras 9
<i>Squalius peloponensis</i> /Miras	37° 17' 34.2" N, 21° 42' 7.86" E	KY070416	KY070485 (a);	KY070567(a);	Miras 10

(NMP PV6 G219)			KY070486 (b)	KY070568 (b)	
<i>Squalius peloponensis</i> /Miras (NMP PV6 G220)	37° 17' 34.2" N, 21° 42' 7.86" E	KY070417	KY070487 (a); KY070488 (b)	KY070569 (a); KY070570 (b)	Miras 11
<i>Squalius peloponensis</i> /Miras (NMP PV6 G222)	37° 17' 34.2" N, 21° 42' 7.86" E	KY070418	KY070489 (a); KY070490 (b)	KY070571 (a); KY070572 (b)	Miras 12
<i>Squalius peloponensis</i> /Miras (MNCN_ ICTIO 94.744)	37° 00' 00.0" N, 21° 44' 23.5" E	KY070404	KY070461 (a); KY070462 (b)	KY070507 (a); KY070508 (b)	Miras 13
<i>Squalius peloponensis</i> /Miras (MNCN_ ICTIO 94.745)	37° 00' 00.0" N, 21° 44' 23.5" E	KY070405	KY070463 (a); KY070464 (b)	KY070509 (a); KY070510 (b)	Miras 14
<i>Squalius peloponensis</i> /Miras (MNCN_ ICTIO 94.746)	37° 00' 00.0" N, 21° 44' 23.5" E	KY070406	KY070465 (a); KY070466 (b)	KY070511 (a); KY070512 (b)	Miras 15
<i>Squalius peloponensis</i> /Pamissos (NMP PV6 F1514)	37° 9'25.21" N, 21°58'43.16" E	KY070387	-	-	Pamissos 1
<i>Squalius peloponensis</i> /Pamissos (NMP PV6 F1515)	37° 9'25.21" N, 21°58'43.16" E	KY070388	-	-	Pamissos 2
<i>Squalius peloponensis</i> /Pamissos (NMP PV6 F1516)	37° 9'25.21" N, 21°58'43.16" E	KY070389	-	-	Pamissos 3
<i>Squalius peloponensis</i> /Pamissos (NMP PV6 F1517)	37° 9'25.21" N, 21°58'43.16" E	KY070390	-	-	Pamissos 4
<i>Squalius peloponensis</i> /Pamissos (NMP PV6 F1518)	37° 9'25.21" N, 21°58'43.16" E	KY070391	-	-	Pamissos 5
<i>Squalius peloponensis</i> /Pamissos (NMP PV6 F1519)	37° 9'25.21" N, 21°58'43.16" E	KY070392	KY070437 (a); KY070438 (b)	KY070529 (a); KY070530 (b)	Pamissos 6
<i>Squalius peloponensis</i> /Pamissos (NMP PV6 F1520)	37° 9'25.21" N, 21°58'43.16" E	KY070393	KY070439 (a); KY070440 (b)	KY070531 (a); KY070532 (b)	Pamissos 7
<i>Squalius peloponensis</i> /Pamissos (NMP PV6 F1521)	37° 9'25.21" N, 21°58'43.16" E	KY070394	KY070441 (a); KY070442 (b)	-	Pamissos 8
<i>Squalius peloponensis</i> /Pamissos (NMP PV6 F1522)	37° 9'25.21" N, 21°58'43.16" E	KY070395	KY070443 (a); KY070444 (b)	KY070533 (a); KY070534 (b)	Pamissos 9

<i>Squalius peloponensis</i> /Pamissos (NMP PV6 F1523)	37° 9'25.21" N, 21°58'43.16" E	KY070396	KY070445 (a); KY070446 (b)	KY070535 (a); KY070536 (b)	Pamissos 10
<i>Squalius peloponensis</i> /Pamissos (NMP PV6 F1524)	37° 9'25.21" N, 21°58'43.16" E	KY070397	KY070447 (a); KY070448 (b)	KY070535 (a); KY070536 (b)	Pamissos 11
<i>Squalius peloponensis</i> /Pamissos (NMP PV6 F1525)	37° 9'25.21" N, 21°58'43.16" E	KY070398	KY070449 (a); KY070450 (b)	KY070537 (a); KY070538 (b)	Pamissos 12
<i>Squalius peloponensis</i> /Pamissos (NMP PV6 F1526)	37° 9'25.21" N, 21°58'43.16" E	KY070399	KY070451 (a); KY070452 (b)	KY070539 (a); KY070540 (b)	Pamissos 13
<i>Squalius peloponensis</i> /Pamissos (NMP PV6 F1527)	37° 9'25.21" N, 21°58'43.16" E	KY070400	KY070433 (a); KY070454 (b)	KY070541 (a); KY070542 (b)	Pamissos 14
<i>Squalius peloponensis</i> /Pamissos (NMP PV6 F1541)	37° 9'25.21" N, 21°58'43.16" E	KY070401	KY070455 (a); KY070456 (b)	KY070543 (a); KY070544 (b)	Pamissos 15
<i>Squalius peloponensis</i> /Pamissos (G244)	37°15'17.39"N, 21°53'45.15" E	KY070402	KY070457 (a); KY070458 (b)	KY070545 (a); KY070546 (b)	Pamissos 16
<i>Squalius peloponensis</i> /Pamissos (G245)	37°15'17.39"N, 21°53'45.15" E	KY070403	KY070459 (a); KY070460 (b)	KY070573 (a); KY070574 (b)	Pamissos 17

4. Molecular clock

To establish a temporal framework for the introgression event, we estimated divergence time among the populations of *Squalius keadicus* (Evrotas R.) and the non-introgressed and introgressed populations of *S. peloponensis* using an uncorrelated relaxed molecular clock based on the *MT-CYB* gene implemented in BEAST v.2.0. (Bouckaert *et al.*, 2014). To incorporate a speciation-model tree prior into the analyses, the birth-death model with incomplete taxon sampling (Stadler *et al.*, 2009), we used one specimen per species to estimate divergence time. For *S. peloponensis*, we used a specimen from the Alfios River (non-introgressed population) and one from the Miras River (introgressed population). As in the phylogenetic analysis, several species of the genus *Petroleuciscus* were used as outgroup. We calibrated the molecular clock using two fossil species of the genus *Squalius* (formerly within the genus *Leuciscus*) as lognormal prior: *Leuciscus antunesi* Gaudant 1977, found in a Portuguese deposit of the Middle Miocene period (13.5-14.5 Ma; Gaudant, 1977; Cabrera y Gaudant, 1985) and *Leuciscus aff. cephalus* from near the Aliakmon River in Greece (Lava 2 deposit) from Upper Miocene (6.56 Ma) (Böhme e Ilg, 2003). The calibration point based on *L. antunesi* was allocated in the stem group of the Iberian *Squalius* species belonging to the Mediterranean lineage, whereas the age of the fossil *Leuciscus aff. cephalus* was placed in the crown of the current *Squalius cephalus* group. The MCMC analyses were run for 100 million generations, with parameters logged every 10000 generations. The remaining parameters were default parameters of the software. Convergence of parameters was evaluated using Tracer v.1.6 (Rambaut y Drummond, 2007) and results were summarized in TreeAnnotator v.2.0 (Bouckaert *et al.*, 2014).

5. Morphological analyses

Morphological analyses included specimens from the Alfios, Miras, Pamissos and Evrotas basins. For each specimen, 23 morphometric and five meristic variables were recorded. All morphometric measurements were performed using digital callipers. Measurements and counts followed (Doadrio *et al.*, 2007) and included standard length (SL), head length (HL), eye diameter (ED), interorbital width (IW), preorbital length (PrOL), postorbital length (PosOL), predorsal distance (PrDD), prepectoral distance (PrPD), preventral distance (PrPV), pre-anal distance (PrPA), caudal peduncle length (CPL), anal peduncle length (APL), pectoral-ventral length (PVL), ventral-anal length (VAL), body depth (BD), body least depth (BLD), dorsal fin height (DFH), dorsal fin length (DFL), anal fin height (AFH), anal fin length (AFL), pectoral fin length (PFL), ventral fin length (VFL), caudal fin length (CFL), lateral line scales (LLS), upper transverse scales (UTS), lower transverse scales (LTS), dorsal fin rays (D) and anal fin rays (A) (Fig. 2).

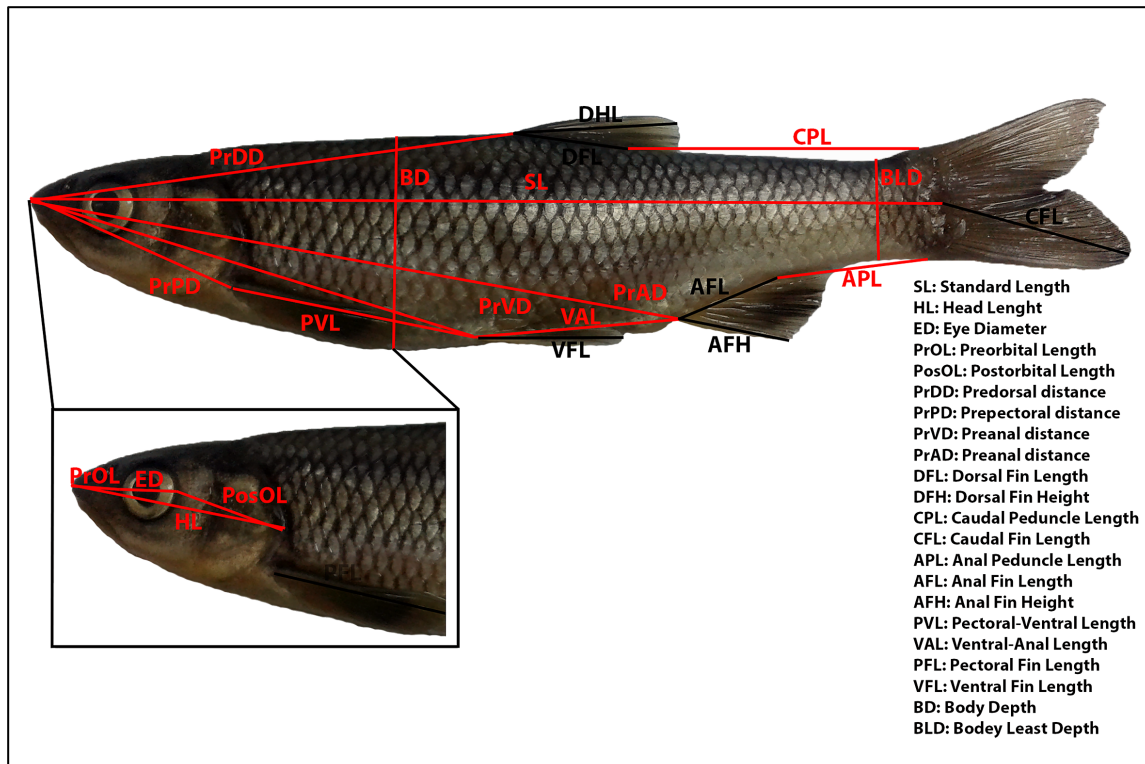


Fig. 2 Morphometric variables measured in the *Squalius* populations analysed.

After constructing the measurement matrix, Burnaby's method was used to correct for size effect in morphometric variables (Burnaby, 1966; Cunha *et al.*, 2004). A canonical variate analysis (CVA) was conducted to identify the variables most explanatory of the morphological variation among the analysed populations with the aim to maximize differences among populations. Non-parametric Kruskal–Wallis and Mann–Whitney *post hoc* comparisons were used to test for significant differences of morphological variables among populations. All analyses were performed using the Burnaby's corrected matrix. Statistical analyses were carried out in PAST v.3.12 software (Hammer *et al.*, 2001).

RESULTS

1. Molecular analyses

The result of mitochondrial phylogenetic analyses is shown in Fig. 3. The *Squalius* populations from the Pamissos and Miras basins clustered together, forming a monophyletic clade, sister to *S. keadicus*. These two clades (*S. keadicus* and Pamissos/Miras *S. peloponensis*) were highly supported as monophyletic in the BI and ML analyses. Nuclear genes yielded tree topology inconsistent with that of the mitochondrial marker, with the Pamissos and Miras populations showing a highly supported cluster with *S. peloponensis* from the Alfios Basin, forming a

monophyletic clade (Figs 4, 5, and 6). All phased alleles for the nuclear genes *RAG1* and *S7* from the Miras and Pamissos populations were nested within the clade of *S. peloponensis* and none of them was grouped with *S. keadicus*. No evidence of genetic structure was observed in the nuclear clade comprising the Alfios, Miras, and Pamissos populations.

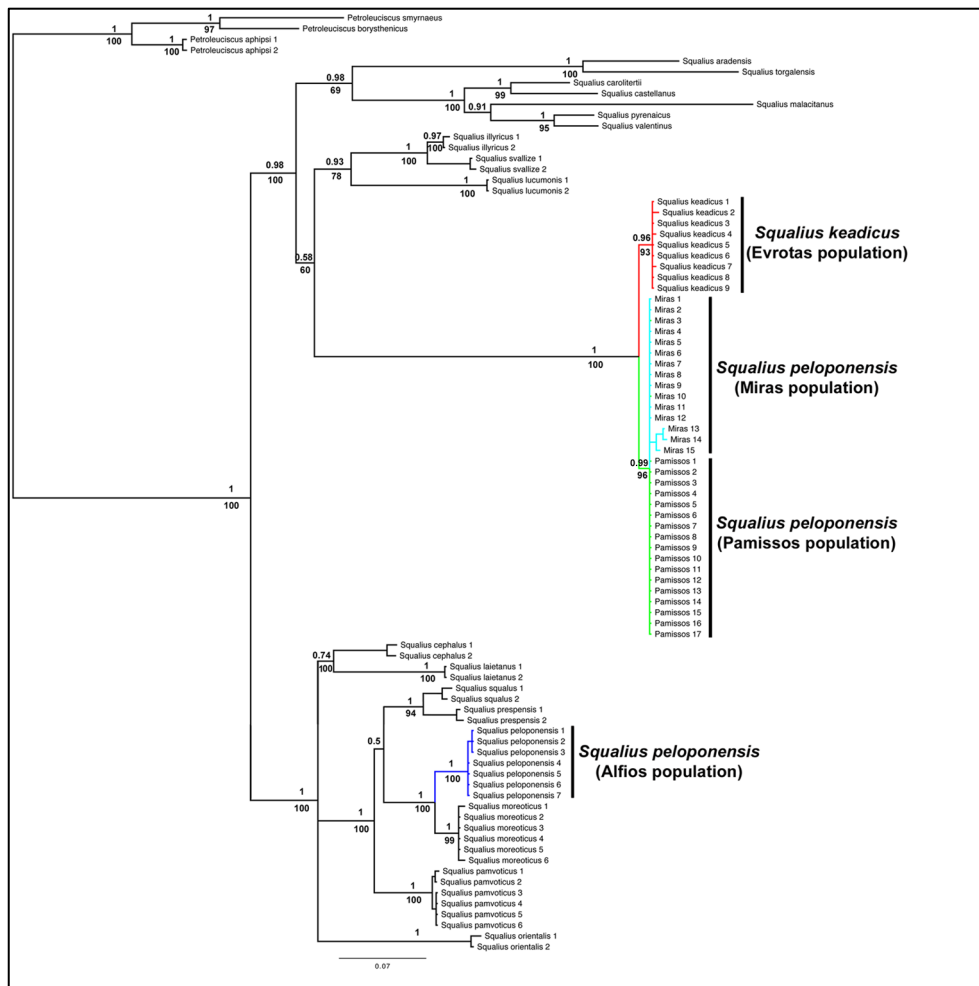


Fig. 3 Phylogenetic tree based on the *MT-CYB* gene. Bootstrap values for ML analysis below branches. Posterior probability values for Bayesian inference above branches.

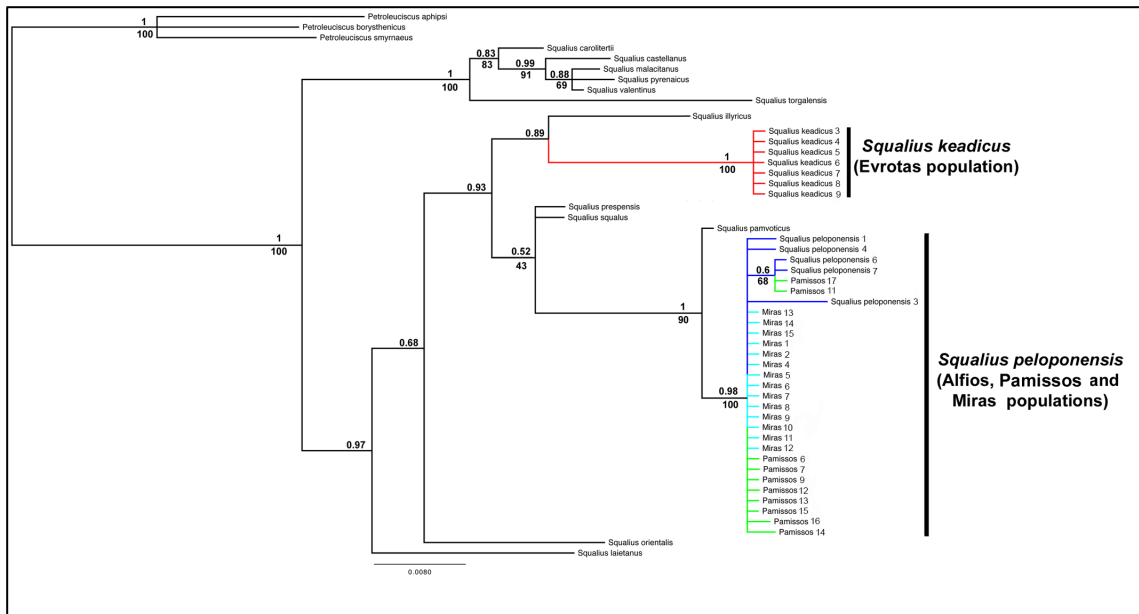


Fig. 4 Phylogenetic tree based on the nuclear S7 gene. Bootstrap values for ML analysis below branches. Posterior probability values for Bayesian inference above branches.

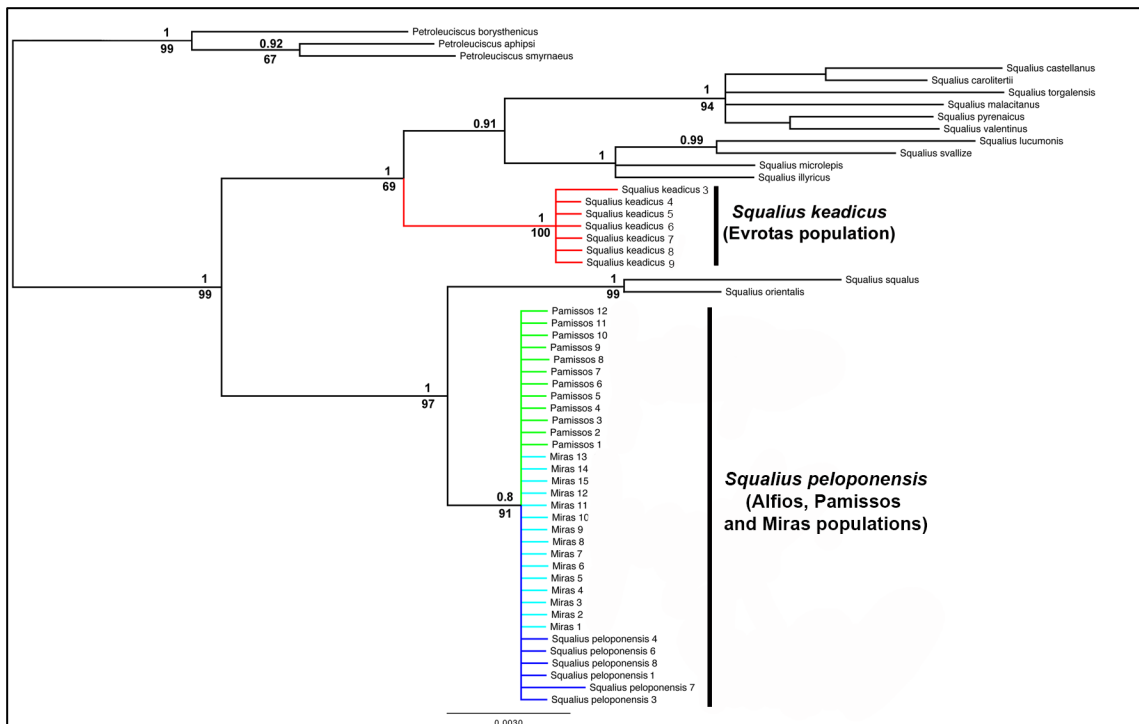


Fig. 5 Phylogenetic tree based on the nuclear RAG1 gene. Bootstrap values for ML analysis below branches. Posterior probability values for Bayesian inference above branches.

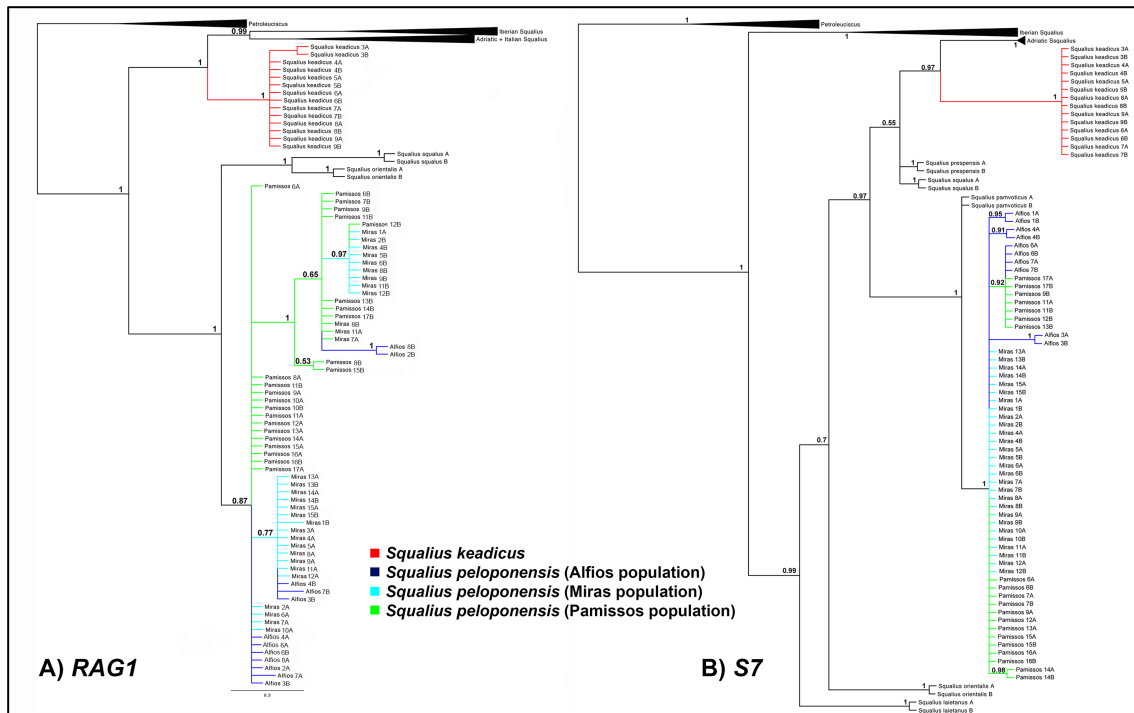


Fig. 6 Phylogenetic tree based on the phased nuclear *RAG1* (A) and *S7* (B) genes. Posterior probability values for Bayesian inference above branches. Blue = Alfios (*S. peloponensis*); Red = Evrotas (*S. keadicus*); Green = *S. keadicus*-like mitochondrial (Miras and Pamissos populations).

Uncorrected-*p* genetic distances between *S. keadicus* and the introgressed populations (Miras and Pamissos) for the three analysed genes are presented in Table 2. S3 Table shows diagnostic autapomorphies characteristic of each population. For the *MT-CYB*, 115 autapomorphies (21 transversions) were recognized in the Alfios Basin population relative to the three other populations, five (1 transversion) were observed in *S. keadicus* and three in the introgressed populations, one of them present only in the Pamissos river system. For the *RAG1* gene, 13 autapomorphies (4 transversions) were found in *S. keadicus* in relation to the three populations recognized as *S. peloponensis*. For the *S7* gene, 23 autapomorphies (9 transversions) appeared in *S. keadicus*, as well as three insertions and three deletions not found in the other three populations.

Estimated divergence times of the mitochondrial lineage of *Squalius keadicus* from that of the introgressed populations of *S. peloponensis* was approximately 0.9 Ma (CI 95% HPD: 0.3-2.2 Ma; Fig. 7). The divergence of *S. peloponensis* from closely related *Squalius* spp occurred in Middle-Upper Pliocene. However, this phylogenetic relationship was not highly supported, probably as a consequence of an incomplete sampling of the Euroasiatic lineage of this genus *sensu* Sanjur *et al.*, 2003 [see also (Kottelat and Freyhof, 2007)].

Table 2. Uncorrected-*p* genetic distances (%) between (below diagonal) and within (in diagonal) populations for the three genes analysed.

<i>MT-CYB</i>				
	Evrotas	Pamissos	Miras	Alfios
Evrotas	0.01			
Pamissos	0.6	0.00		
Miras	0.6	0.1	0.01	
Alfios	11.0	10.7	10.8	0.01
<i>S7</i>				
	Evrotas	Pamissos	Miras	Alfios
Evrotas	0.0			
Pamissos	2.8	0.1		
Miras	2.7	0.1	0.0	
Alfios	2.9	0.2	0.2	0.3
<i>RAG1</i>				
	Evrotas	Pamissos	Miras	Alfios
Evrotas	0.0			
Pamissos	1.0	0.1		
Miras	1.0	0.2	0.1	
Alfios	1.0	0.1	0.2	0.1

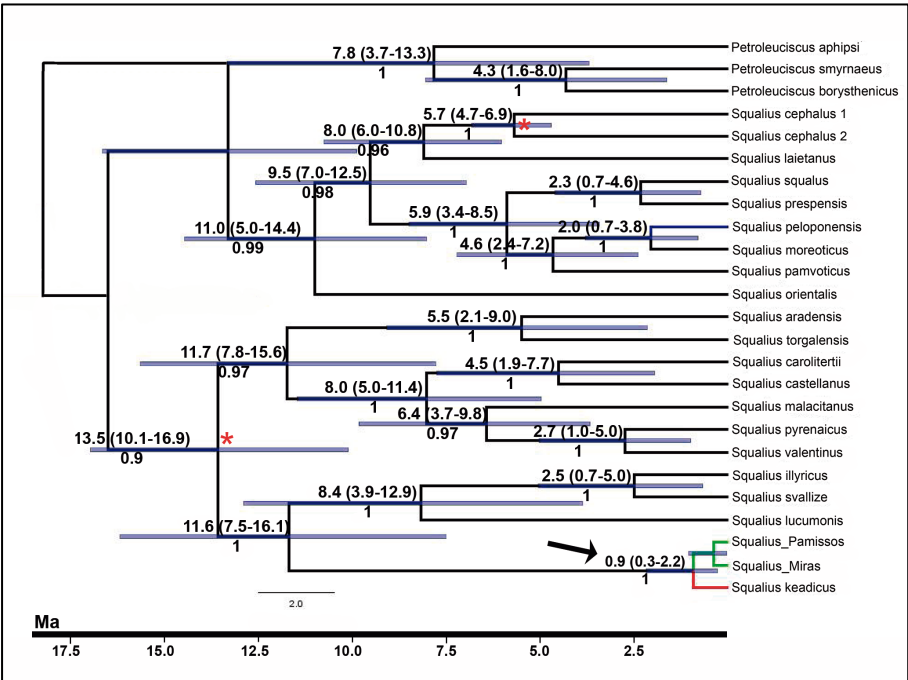


Fig. 7 Chronogram obtained for the *MT-CYB* gene for several *Squalius* species. Divergence time estimates and their HPD 95% confidence intervals, in Ma, for the main cladogenetic events between *Squalius* species above branches. Posterior probability values of Bayesian inference below branches (** indicates posterior probability = 1). Black arrow highlight the divergence time estimated for the separation of mitochondrial lineages of *S. keadicus* and the introgressed populations. Red asterisk indicates fossil ages used to calibrate the molecular clock.

2. Morphological analyses

Kruskall–Wallis and Mann–Whitney *post hoc* analyses demonstrated significant differences among the four analysed *Squalius* populations in morphometric and meristic variables (Table 3; Fig. 8). Canonical variate analysis (CVA) also supported morphological differences, showing four isolated groups with partial overlap between the populations from the Miras and Pamissos basins and between each of these populations with both the Evrotas (*S. keadicus*) and Alfios (*S. peloponensis*) populations. As expected, overlap between *S. keadicus* and the Alfios *S. peloponensis* population was not observed (Fig. 9). The most contributory morphometric variables to the ordination in the CVA were AFL, BD, and DFL (Table 4).

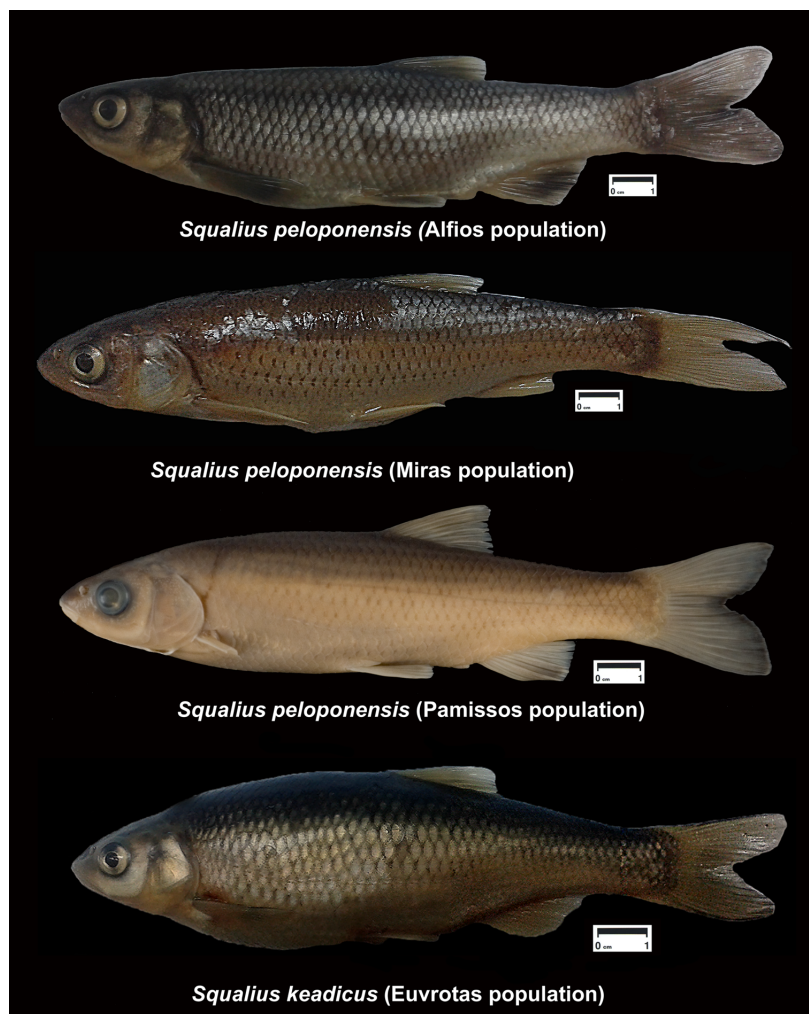


Fig. 8 Specimens of the four analysed populations. Voucher specimen numbers: *Squalius peloponensis* from Alfios River (MNCN_ICTIO 94.982); *Squalius peloponensis* from Pamissos River (NMP PV6 F1227); *Squalius peloponensis* from Miras River (MNCN_ICTIO 94.747); *Squalius keadicus* from Evrotas River (MNCN_ICTIO 123.875).

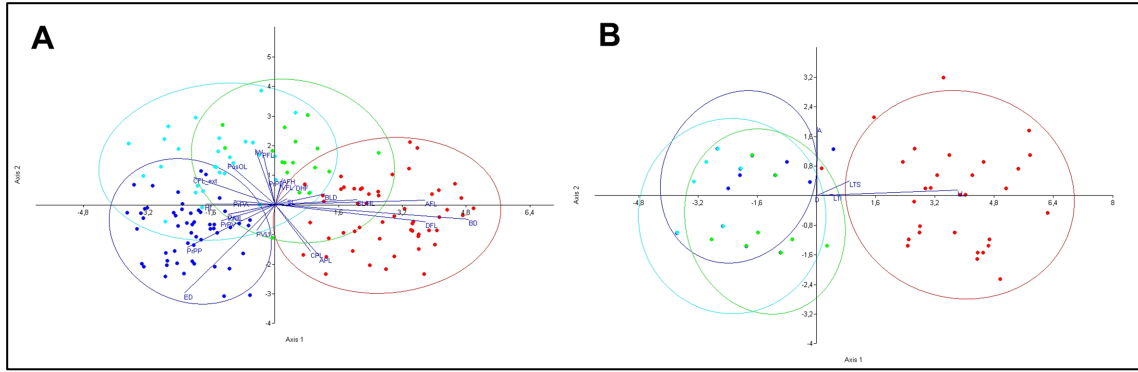


Fig. 9 Canonical Variate Analysis of the four *Squalius* populations. A) Morphometric measurements. B) Meristic measurements. Red: Evrotas population (*S. keadicus*), blue: Alfios population (*S. peloponensis*); light blue: Miras population (*S. peloponensis*); green: Pamissos population (*S. peloponensis*).

Table 3. Kruskal-Wallis test and non-parametric Mann-Whitney pairwise *post hoc* comparisons for all populations. Ev = Evrotas, Al = Alfios, Pa = Pamissos, Mi = Miras, Significant differences $p < 0.01$. Acronyms of variables are defined in the Material and Method section.

Variable	H	P-value	Significant Mann-Whitney pairwise comparisons
SL	25.57	<0.0001	Al-Ev; Al-Pa
HL	78.95	<0.0001	Al-Ev; Al-Pa; Ev-Mi; Ev-Pa
ED	63.72	<0.0001	Al-Ev; Al-Pa; Al-Mi
IW	56.48	<0.0001	Al-Ev; Al-Pa; Al-Mi; Ev-Pa;
PrOL	19.51	0.0002	Al-Ev
PosOL	30.77	<0.0001	Al-Ev; Pa-Ev; Mi-Ev
PrPD	11.38	0.0099	-
PrPP	92.4	<0.0001	Al-Ev; Al-Pa; Al-Mi; Ev-Mi; Pa-Mi
PrPV	86.22	<0.0001	Al-Ev; Al-Pa; Al-Mi; Ev-Mi; Ev-Pa
PrPA	59.86	<0.0001	Ev-Al; Ev-Pa; Ev-Mi
DHF	13.39	0.0039	Al-Ev
DFL	84.98	<0.0001	Ev-Al; Ev-Pa; Ev-Mi; Al-Pa
CPL	16.06	0.0011	Ev-Mi
CFL	46.87	<0.0001	Ev-Al; Ev-Mi
BLD	16.6	0.0009	Ev-Al; Ev-Mi
BD	100.8	<0.0001	Ev-Al; Ev-Pa; Ev-Mi; Al-Pa
APL	21.72	<0.0001	Ev-Mi
AFL	56.89	<0.0001	Al-Ev; Al-Pa; Ev-Mi
AFH	7.91	0.0479	-
VAL	17.37	0.0006	Al-Ev
PVL	4.794	0.1875	-
VFL	10.42	0.0154	Al-Mi
PFL	11.64	0.0087	Mi-Al; Mi-Ev
LL	106.6	<0.0001	Ev-Al; Ev-Pa; Ev-Mi
UTS	74.29	<0.0001	Ev-Al; Ev-Pa; Ev-Mi; Al-Mi
LTS	40.74	<0.0001	Ev-Al; Ev-Pa; Ev-Mi
D	0.5093	0.9169	-
A	19.1	0.0002	Al-Pa; Al-Mi

Table 4. Canonical Variate Analysis values for the first three canonical axes of analyses of morphometric and meristic variables. Acronyms of variables are defined in the Material and Method section.

Variable	Axis 1	Axis 2	Axis 3
Eigenvalue 1 = 4.861 (77.76 %)			
Eigenvalue 2 = 0.9997 (15.99 %)			
SL	0.0009	0.0008	0.0028
HL	-0.0055	0.0003	0.0002
ED	-0.0069	-0.0094	0.0031
IW	-0.0016	0.0060	-0.0002
PrOL	-0.0038	-0.0014	-0.0049
PosOL	-0.0037	0.0048	0.0015
PrPD	-0.0005	0.0029	0.0003
PrPP	-0.0069	-0.0043	-0.0046
PrPV	-0.0042	-0.0015	0.0013
PrPA	-0.0032	0.0006	0.0018
DFH	0.0017	0.0029	-0.0030
DFL	0.0119	-0.0015	-0.0075
CPL	0.0027	-0.0051	0.0019
CFL	-0.0068	0.0021	-0.0028
BLD	0.0040	0.0008	0.0029
BD	0.0153	-0.0009	-0.0042
APL	0.0032	-0.0046	0.0085
AFL	0.0123	-0.0002	0.0070
AFH	0.0006	0.0032	-0.0044
VAL	-0.0037	-0.0012	0.0056
VFL	0.0004	0.0026	-0.0064
PVL	-0.0014	-0.0033	-0.0014
PFL	-0.0009	0.0051	-0.0028
Eigenvalue 1 = 8.176 (97.3 %)			
Eigenvalue 2 = 0.21 (2.5 %)			
LL	0.9484	0.0370	-0.0734
UTS	0.2201	0.0981	0.4004
LTS	0.1131	0.0053	-0.1368
D	-0.0126	-0.0023	0.0229
A	0.0051	0.4609	-0.0231

Meristic variables in the Alfios, Pamissos, and Miras populations overlapped with the Evrotas population (*S. keadicus*) forming a differentiated group in CVA (Fig. 9). Kruskal–Wallis and Mann–Whitney *post hoc* comparisons were significant for LL, LTS, LTI, and A, reaching higher values in *S. keadicus* relative to the other three populations (Fig. 10).

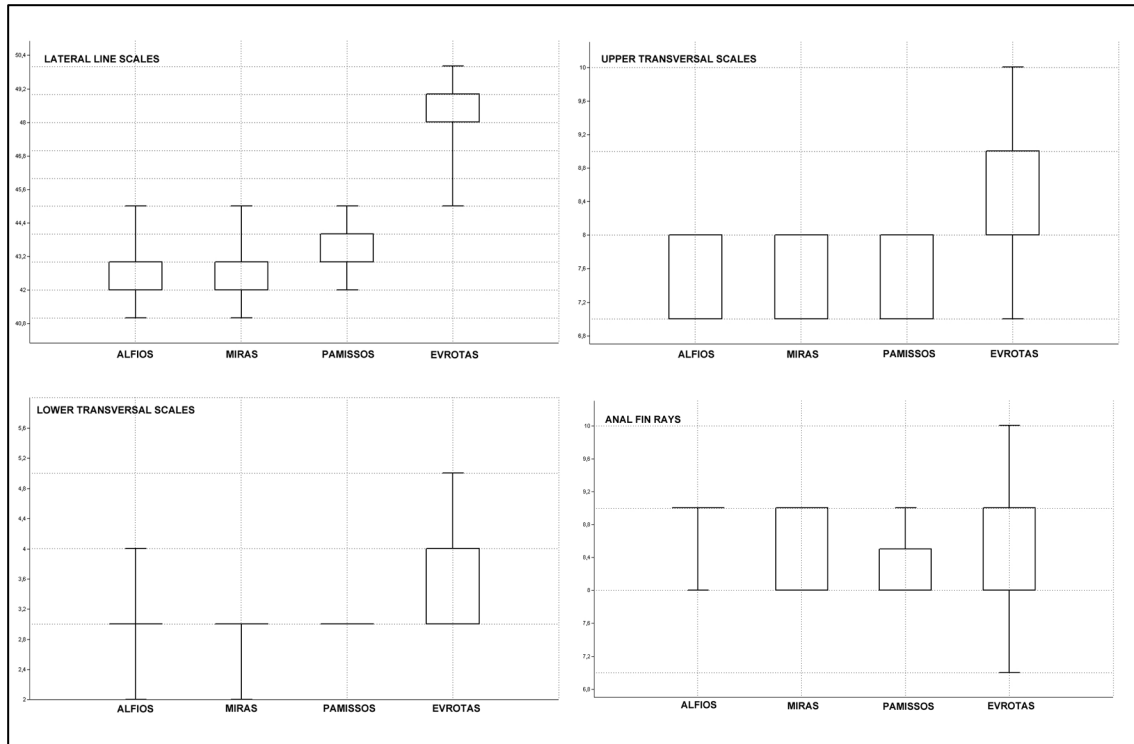


Fig. 10 Box-plot of meristic variables. The 25-75 percent quartiles are drawn inside the box. Short horizontal lines represent minimal and maximal values.

The Evrotas population showed greater BD and longer APL, DFL and AFL fins than the other three populations. It also exhibited significantly shorter HL relative to SL than in other populations (Table 5). The Pamissos population exhibited the significantly smaller ED relative to HL and SL than seen in the other three populations. The largest ED relative to the standard length was found in the *S. peloponensis* from Alfios (Table 5). Although the raw data matrix was corrected by Burnaby's method, differences in ED may be attributed to allometric growth rather than to real population differences. The Alfios population had the narrowest head in relation to HL (Table 5), a ratio significantly different from the Evrotas and Miras populations, but not from that of the Pamissos.

Table 5. Significant proportions (%) relative to standard length (SL) or head length (HL) of several morphometric variables differentiating the four analysed populations.

Variable	Population	% Relative to SL or to HL
BD relative to SL	Evrotas	25.9 (24.0-28.7)
	Alfios	21.6 (21.2-21.5)
	Miras	22.4 (21.1-22.7)
	Pamissos	22.5 (23.0-24.7)
DFL relative to SL	Evrotas	12.6 (12.0-13.4)
	Alfios	11.1 (10.4-11.4)
	Miras	11.3 (10.5-11.4)
	Pamissos	11.5 (10.8-12.0)
AFL relative to SL	Evrotas	11.4 (8.9-12.6)
	Alfios	10.4 (8.7-12.3)
	Miras	10.2 (9.3-10.0)
	Pamissos	11.5 (10.5-12.7)
APL relative to SL	Evrotas	21.4 (21.0-21.7)
	Alfios	20.8 (20.9-21.1)
	Miras	20.2 (18.0-18.9)
	Pamissos	20.5 (19.9-19.9)
HL relative to SL	Evrotas	23.5 (24.4-24.6)
	Alfios	25.1 (24.9-26.3)
	Miras	24.9 (23.5-26.8)
	Pamissos	23.8 (23.6-26.0)
ED relative to HL	Evrotas	24.6 (24.0-25.5)
	Alfios	27.0 (24.8-28.9)
	Miras	27.0 (24.4-29.5)
	Pamissos	24.6 (24.0-25.5)
ED relative to SL	Evrotas	6.3 (6.0-7.2)
	Alfios	6.8 (6.2-7.6)
	Miras	6.5 (5.5-6.6)
	Pamissos	5.8 (5.9-6.6)
IW relative to HL	Evrotas	42.5 (39.2-40.2)
	Alfios	36.0 (35.4-39.6)
	Miras	42.0 (39.7-41.6)
	Pamissos	41.6 (41.2-41.9)

DISCUSSION

In this study we tested through morphological and molecular analyses the hypothesis of mitochondrial introgression between two species of the genus *Squalius* from Greece, *S. keadicus* and *S. peloponensis*, formulated on the basis of the discordance between mitochondrial (MT-CYB; (Durand *et al.*, 2000)) and nuclear (allozymes; (Doadrio *et al.*, 1998)) genomes found in one population from the Greek Peloponnese (from Miras Basin).

The populations from the Miras and Pamissos basins were morphologically similar to *S. peloponensis* from Alfios Basin, as expected based on sharing of the nuclear genome and the lack of diagnostic nuclear autapomorphies among the three populations (S3 Table), as well as on meristic characters, which fall within the range of *S. peloponensis* (Fig. 10). However, conserved meristic measurements are observed in *Squalius* species of the Euroasiatic lineage *sensu* Sanjur *et al.*, (2003); see also Kottelat and Freyhof, 2007). Although the populations from Miras and Pamissos were recognized as *S. peloponensis* significant differences between these two populations were found in some morphometric measurements, as well as between *S. peloponensis* and *S. keadicus* (Tables 3 and 5; Fig. 9).

Whereas the Miras and Pamissos populations showed the nuclear genome and morphology typical of *Squalius peloponensis*, their mitochondrial genome clustered with *S. keadicus* (Evrotas population), but not with the *S. peloponensis* (Alfios population). This mitonuclear discordance supports the idea of mitochondrial introgression of *S. keadicus* in the Miras and Pamissos populations. Sharing of similar morphology and the nuclear genome allowed us to reject the hypothesis of Miras and Pamissos populations being hybrids of *S. keadicus* and *S. peloponensis*, since they should show similarity to both hypothetical parent species in the nuclear genome as well as in the mitochondrial genome. Therefore, the hypothesis of mitochondrial DNA introgression cannot be rejected based on our results. Hybridization is common in fish, and the boundaries between species may be obscure. In cyprinids, hybridization is a common evolutionary process [26, 66-69]. In our analyses, the two nuclear markers were *S. peloponensis*-like in introgressed populations, and we did not find evidence of recombination between species (PHI=1 for both nuclear markers), as demonstrated by the phylogenetic relationships of phased alleles (Fig 6), and low (0.1-0.2%) genetic distance between populations in both nuclear markers. Hence, we accept the hypothesis of mitochondrial introgression and reject the hypothesis of hybrid speciation.

All analysed specimens from Miras and Pamissos were mitochondrially *S. keadicus*-like and nuclearly *S. peloponensis*-like, suggesting unidirectional complete mitochondrial

introgression (mitochondrial capture) of *S. keadicus* to *S. peloponensis* (Durand *et al.*, 2000). Nonetheless, they formed an independent highly supported phylogenetic clade, sister to *S. keadicus* from the Evrotas population (Fig. 3). These two highly supported reciprocal monophyletic groups found in the mitochondrial tree within the phylogenetic lineage of *S. keadicus* indicate that the introgression event was ancient, but it is not clear whether divergence of these populations, allowing mitochondrial DNA to sort, occurred before or after the introgression event. Ancient mitochondrial captures have also been described for reptiles (Leaché and McGuire, 2006; Barbanera *et al.*, 2009), insects (Linnen and Farrell, 2007), and fishes (Nevado *et al.*, 2009; Tang *et al.*, 2012). Since these four *Squalius* populations currently occur in allopatry, a geographical barrier may be considered the primary source of their morphological differences possibly through a process of incipient speciation.

Isolation of populations and incipient speciation in sympatry can be attributed to sexual selection (Alexander and Breden, 2004; Meyer *et al.*, 2006), divergent selection, and phenotypic plasticity (Magalhaes *et al.*, 2009), or segregation of ecological niches (Knudsen *et al.*, 2006; Nosil *et al.*, 2009). The morphological and molecular differentiation found among the analysed populations of *Squalius* from western Peloponnese probably occurred after their allopatric separation linked to the formation of the hydrological network in the region during the Upper Pliocene-Pleistocene periods (Athanasas and Fountoulis, 2013; Fountoulis *et al.*, 2014). Incipient speciation or occurrence of differentiated lineages during the Pleistocene as a consequence of geographic barriers has been suggested for several taxa (Perdices *et al.*, 2003; Vila *et al.*, 2005; Bohlen *et al.*, 2006). Examples of isolation-driven divergence to new lineages after ancient hybridization have been reported (Bryson *et al.*, 2010, 2014). This is demonstrated here by the lack of shared mitochondrial haplotypes among populations in the rivers of the western Peloponnese. The sharing of the nuclear genome among the Alfios, Miras, and Pamissos populations ascribed to *S. peloponensis*, clustering in the same clade, also supports the scenario of incipient speciation.

Several biogeographic scenarios may be proposed to explain the mitochondrial introgression event and the current molecular configuration of the four analysed populations (Fig. 11): A) The lineage of *Squalius peloponensis* inhabited the Alfios, Miras, and Pamissos basins, and at some time the lineage of *S. keadicus* invaded the Miras and Pamissos basins. The introgression took place, and subsequently mitogenomes of the mitochondrial lineages of *S. keadicus* from the Evrotas basin and the introgressed populations of the Miras and Pamissos diverged. B) The lineage of *Squalius keadicus* inhabited the Evrotas, Miras, and Pamissos basins, and at some time the lineage of *S. peloponensis* invaded the Miras and Pamissos basins. The introgression event took place, mitogenomes of the lineages of *S. keadicus* from the Evrotas

basins and the introgressed Miras and Pamissos populations subsequently diverged. C) Miras and Pamissos basins were inhabited by the lineage of *Squalius keadicus* and the mitogenome of these populations diverged from that of the Evrotas population. The lineage of *S. peloponensis* then invaded the Miras and Pamissos basins and introgression occurred. D) A case similar to scenario C is that in which it was not the mitochondrial lineage of *Squalius keadicus* that inhabited the Miras and Pamissos basins prior to the invasion by the lineage of *S. peloponensis*, but a phylogenetically closely related species.

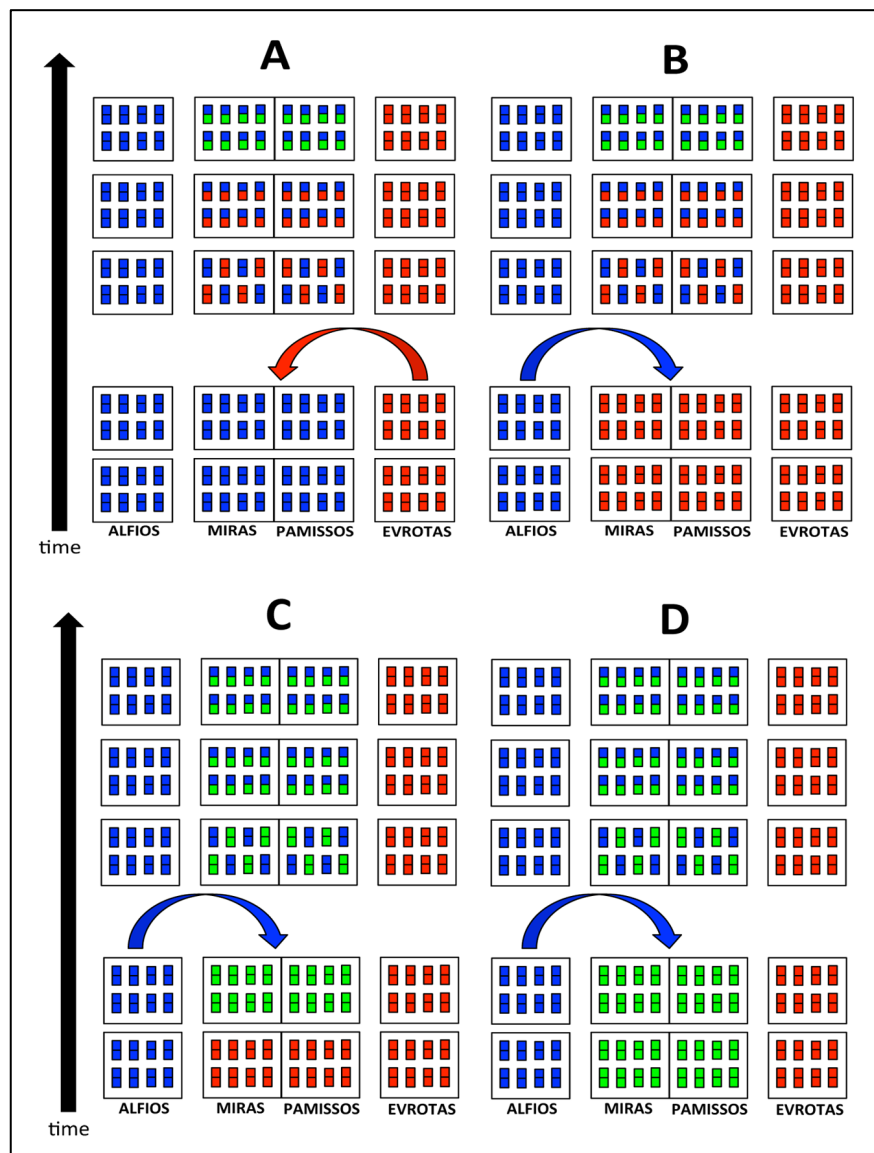


Fig. 11 Hypothetical biogeographical scenarios explaining the mitochondrial capture event from *Squalius keadicus* to *Squalius peloponensis*. Each rectangle symbolizes a hypothetical individual within the analysed population. The two squares within each rectangle symbolize the mitochondrial and the nuclear genome of this hypothetical individual in the population. Colours indicate populations: Blue = *S. peloponensis* lineage; Red = *S. keadicus* lineage; Green = *S. keadicus*-like mitochondrial lineage (Miras and Pamissos populations).

The main difference in these biogeographical scenarios is the timing of the introgression event as before (scenarios A and B) or after (scenarios C and D) the divergence of the *S. keadicus* from the *S. keadicus*-like mitogenome. The common factor among the hypothetical scenarios is the biogeographical discordance between mitochondrial and nuclear DNA [4]. Most patterns of biogeographic mitonuclear discordance are characterized by previously isolated lineages that came into contact and generated hybrid zones in which gene flow patterns in mitochondrial and nuclear DNA differed (Toews and Brelsford, 2012). This hypothesis of secondary contact is consistent with the hypothetical biogeographical scenarios that we propose, which differ only whether the invader species is *S. keadicus* (A) or *S. peloponensis* (B, C, and D). Differences in effective population size of the resident and the invader species could be responsible for the unidirectionality of the introgression event found in the present study (Funk and Omland, 2003; Toews and Brelsford, 2012; Rheindt and Edwards, 2011). In the initial stages of an invasion event, a small number of invader specimens are expected in comparison to the resident population, and simulated studies have revealed unidirectional introgression of neutral loci from the resident to the invader population (Currat *et al.*, 2008). This would imply that *S. keadicus* (B, C) or *S. keadicus*-like (D) was the resident species and *S. peloponensis* the invader. Nevertheless, if the invader becomes more numerous as invasion continues, the direction of the introgression will depend of the demographic balance of the resident and invader populations (Ludwig *et al.*, 2003; Currat *et al.*, 2008). Mitochondrial DNA introgression as a consequence of extreme differences in population size has been described in other freshwater fish species (Knudsen *et al.*, 2006). In our hypothetical scenarios, *S. keadicus* or an extinct species phylogenetically similar to *S. keadicus*, regardless of whether it was the resident or the invader species, must have had a population large enough to erode the mitochondrial genome of *S. peloponensis* to the point of elimination. Another possible source of mitonuclear discordance may be sex-biased reproduction or sex-biased offspring production during the hybridization event (Toews and Brelsford, 2012). Empirical experiments are necessary to assess this hypothesis.

Differences in distribution range of two species involved in an introgression event when a hybrid zone is formed explain asymmetric mitochondrial introgression, associated with differences in population size and competition between their mitochondrial genomes (Toews and Brelsford, 2012). It is not possible to estimate the range of *S. keadicus* or the size of the hybrid zone in the Miras and Pamissos basins, because *S. keadicus*, or the *S. keadicus*-like species, is currently extinct in this region and no fossil evidence exists, but it is clear that the distribution range of the studied species overlapped in the past.

Although the mitochondrial genome is generally considered to be neutral (McKay and Zink, 2010), positive mitochondrial selection has been described for this marker (Ballard and Whitlock,

2004), and is another possible explanation of the observed unidirectional mitochondrial introgression. Thus, the mitochondrial genome with the highest fitness may be introgressed into the other species, regardless of whether it is resident or invader (Rheindt and Edwards, 2011). These selective sweeps have been attributed to climatic adaptation, especially to temperature (Cheviron and Brumfield, 2009). Our study area, however, was restricted to a small geographical region with no significant climatic differences in river systems. Furthermore, the HKA test of selection (Hudson et al., 1987) performed on the *MT-CYB* gene ($p > 0.05$) suggested the potential for neutral evolution in this gene, and a selective sweep is not indicated in the studied species.

The most recent common ancestor of the mitochondrion of the Evrotas *S. keadicus* and the introgressed populations of the Miras and Pamissos basins dated from 0.9 Ma (CI 95% HPD 0.3-2.2 Ma) in the Pleistocene, and introgression may have occurred before or after divergence. Mitochondrial captures during Pleistocene have been described for other freshwater fish species (Tang et al., 2012). The divergence of the mitochondrial clades in the present study may be attributed to Pleistocene tectonic movements related to uplift and river system formation. The paleogeographical history of western Greece and the Peloponnese is complex; the area comprises a network of active faults that constitutes a multifractured neotectonic macrostructure (Fountoulis et al., 2014). Tectonic movements occurred throughout the Upper Pliocene-Pleistocene period and continue to the present causing earthquakes and tsunamis (Kokkalas et al., 2006; Ntzgeretzis et al., 2015). The faults had great influence on the drainage network formation (Fountoulis et al., 2014). The current hydrological basins of the western Peloponnese were formed during the Pleistocene as a consequence of marine or continental sediment infilling when fault systems were activated. Marine sediments found at several hundred metres elevation in some mountainous areas are evidence of vertical tectonic movements throughout this period [(Fountoulis et al., 2014; Athanassas, 2010). These sedimentary events have contributed to configuration of the main fluvial basins of western Peloponnese, which show a radial pattern commonly found in areas of rapid uplifting (Athanassas and Fountoulis, 2013).

The recent paleogeographic events may be responsible for the divergence of mitochondrial lineages of *S. keadicus* and the introgressed populations (mitochondrial *S. keadicus*-like) of *S. peloponensis* as well as morphological differences among the four populations that evolved following the isolation of the hydrological basins (Fig. 11). Complex tectonic movements configuring the hydrological network of western Peloponnese may also have been the source of contact between species and the introgression event, through river piracy.

CONCLUSIONS

Molecular and morphological evidence supports the hypothesis of mitochondrial introgression of *S. keadicus* to *S. peloponensis*. We found unidirectional mitochondrial introgression from *S. keadicus* into populations from two basins of the western Peloponnese, the Miras and Pamissos River systems, but not into that from the Alfios Basin, and a sorting of the mitochondrial genome of introgressed populations from that of *S. keadicus*, likely as a consequence of isolation by biogeographical barriers, leading to a process of incipient speciation. Recent secondary contact among basins is the most plausible hypothesis to explain the mitochondrial introgression.

The identification of mitochondrial capture is crucial to understanding the evolutionary history of living organisms, and adequate molecular and morphological analyses are essential to distinguish among evolutionary processes leading to mitonuclear discordance, such as the case of mitochondrial introgression.

Acknowledgements

We thank I. Karakousis, P. Economidis, P. Garzón, A. Perdices and J.A. Carmona for field assistance. The Lucidus Consultancy helped us with English editing of the manuscript.

REFERENCES

- Alexander HJ, Breden F. 2004. Sexual isolation and extreme morphological divergence in the Cumaná guppy: a possible case of incipient speciation. *Journal of Evolutionary Biology*, 17: 1238-1254.
- Aboim MA, Mavárez J, Bernatchez L, Coelho MM. 2010. Introgressive hybridization between two Iberian endemic cyprinid fish: a comparison between two independent hybrid zones. *Journal of Evolutionary Biology*, 23(4): 817-828.
- Aguilar A, Jones WJ. 2009. Nuclear and mitochondrial diversification in two native California minnows: insights into taxonomic identity and regional Phylogeography. *Molecular Phylogenetics and Evolution*, 51: 373-381.
- Athanassas C. 2010. *Optically stimulated luminescence dating of Late Quaternary marine terraces, sea-level changes and coastal neotectonics in W. Messinia (SW Peloponnese, Greece)*. Unpublished D. Phil. Thesis. Faculty of Geology and Geoenvironment. National and Kapodistrian University of Athens. Greece. (In Greek).
- Athanassas C, Fountoulis I. 2013. Quaternary neotectonic configuration of the southwestern Peloponnese, Greece, based on luminescence ages of marine terraces. *Journal of Earth Sciences*, 24(3): 410-427.
- Ballard JWO, Whitlock MC. The incomplete natural history of mitochondria. *Molecular Ecology*, 13: 729-744.
- Barbanera F, Zuffi MAL, Guerrini M, Gentili A, Tofanelli S, Fasola M, Dini F. 2009. Molecular Phylogeography of the asp viper *Vipera aspis* (Linnaeus, 1758) in Italy: evidence for introgressive hybridization and mitochondrial DNA capture. *Molecular Phylogenetics and Evolution*, 52: 103-114.
- Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu C-H, Xie D, Suchard MA, Rambaut A, Drummond AJ. 2014. BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLoS Computational Biology*, 2014; 10(4): e1003537.
- Böhme, M, Ilg, A. 2003. fosFARbase, www.wahre-staerke.com/ (Accessed April 2016).

- Bohlen J, Perdices A, Doadrio I, Economidis PS. 2006. Vicariance, colonisation, and fast local speciation in Asia Minor and the Balkans as revealed from the phylogeny of spined loaches (Osteichthyes; Cobitidae). *Molecular Phylogenetics and Evolution*, 39: 552-561.
- Broughton RE, Vedala KC, Crowl TM, Ritterhouse LL. 2011. Current and historical hybridization with differential introgression among three species of cyprinid fishes (genus *Cyprinella*). *Genetica*, 139: 699-707.
- Bruen TC, Phillipe H, Bryant D. 2006. A simple and robust statistical test for detecting the presence of recombination. *Genetics*, 172(4): 2665-2681.
- Burnaby TP. 1966. Growth-invariant discriminant functions and generalized distances. *Biometrics*, 22: 96-110.
- Bryson RW Jr, Nieto-Montes de Oca A, Jaeger JR, Riddle BR. 2010. Elucidation of cryptic diversity in a widespread Nearctic treefrog reveals episodes of mitochondrial gene capture as frogs diversified across a dynamic landscape. *Evolution*, 64: 2315-2330.
- Bryson RW Jr, Smith BT, Nieto-Montes de Oca A, García-Vázquez UO, Riddle BR. The role of mitochondrial introgression in illuminating the evolutionary history of Nearctic treefrogs. *Zoological Journal of the Linnean Society*, 172: 103-116.
- Chan KMA, Levin SA. Leaky prezygotic isolation and porous genomes: rapid introgression of maternally inheritance DNA. *Evolution*, 2005; 59: 720-729.
- Cheviron ZA, Brumfield RT. 2009. Migration-selection balance and local adaptation of mitochondrial haplotypes in Rufous-collared sparrows (*Zonotrichia capensis*) along an elevational gradient. *Evolution*, 63: 1593-1605.
- Costedoat C, Pech N, Chappaz R, Gilles A. 2007. Novelities in hybrid zones: Crossroads between population genomic and ecological approaches. *PLoS One*, 2(4): e357.
- Cunha C, Bastir M, Coelho MM, Doadrio I. 2004. Body shape evolution among ploidy levels of the *Squalius alburnoides* hybrid complex (Teleostei, Cyprinidae). *Journal of Evolutionary Biology*, 2004; 22: 718-728.
- Cunha C, Coelho MM, Carmona JA, Doadrio I. 2004. Phylogeographical insights into the origins of the *Squalius alburnoides* complex via multiple hybridization events. *Molecular Ecology*, 13(9): 2807-2817.
- Currat M, Ruedi M, Petit RJ, Excoffier L. The hidden side of invasions: massive introgression by local genes. *Evolution*, 62(8): 1908-1920.
- Cuvier G, Valenciennes A. *Histoire naturelle des poissons. Tome dix-septième. Suite du livre dix-huitième. Cyprinoïdes*. v.17: i-xxiii + 1-497 + 2 pp.; 1844. p. 487-519.
- Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*, 9: 772.
- Degnan JH, Rosenberg NA. 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends in Ecology and Evolution*, 24(6): 332-340.
- Doadrio I, Carmona JA. 1998. Genetic diversity in Greek of the genus *Leuciscus* and its evolutionary and biogeographical implications. *Journal of Fish Biology*, 53(3): 591-613.
- Doadrio I, Perea S, Alonso F. A new species of the genus *Squalius* Bonaparte, 1837 in the Tagus Basin (Central Spain). *Graellsia*, 2007; 63(1): 89-100.
- Durand JD, Ünlu E, Doadrio I, Pipoyan S, Templeton AR. Origin, radiation, dispersion and allopatric hybridization in the chub *Leuciscus cephalus*. *Proceedings of the Royal Society of Biological Science*, 267: 1687-1697.
- Edwards S, Bensch S. 2009. Looking forwards or looking backwards in avian phylogeography? A comment on. *Molecular Ecology*, 18(14): 2930-2933.
- Fountoulis I, Mariolakis I. 2008. Neotectonic folds in the central-western Peloponnese, Greece. *Z. Dtsch. Ges. Geowiss*, 159(3): 485-494.
- Fountoulis I, Mariolakis I, Ladas I. 2014. Quaternary basin sedimentation and geodynamics in SW Peloponnese (Greece) and late stage uplift of Taygetos Mt. *B. Geofis. Teor. Appl.*, 55(2): 303-324.

- Funk DJ, Omland KE. 2003. Species-level paraphyly and polyphyly: Frequency, causes, and consequences, with insights from Animal Mitochondrial DNA. *Annual Review of Ecology and Evolutionary Systematics*, 34: 397-423.
- Freyhof J, Dietmar L, Pitra C, Ludwig A. 2005. Molecules and morphology: Evidence for introgression of mitochondrial DNA in Dalmatian cyprinids. *Molecular Phylogenetics and Evolution*, 37: 347-354.
- Gaudant J. 1977. Contributions à la Paléontologie du Miocène moyen continental du bassin du Tage. II-Observations sur les dents pharyngiennes de Poissons Cyprinidés-Póvoa de Santarém. *Ciencias da Terra*, 3: 129-141.
- Geiger M, Schreiner C, Delmastro GB, Herder F. Combining geometric morphometrics with molecular genetics to investigate a putative hybrid complex: a case study with barbels *Barbus* spp. (Teleostei: Cyprinidae). *Journal of Fish Biology*, 88(3): 1038-1055.
- Good JM, Hird S, Reid N, Demboski JR, Stepan SJ, Martin-Nims TR, Sullivan J. 2008. Ancient hybridization and mitochondrial capture between two species of chipmunks. *Molecular Ecology*, 17: 1313-1327.
- Hammer Ø, Harper DAT, Ryan D. 2001. PAST: Paleontological Statistics Software Package for Education and Data Analysis. *Paleontological Electronica*, 41(1): 9pp.
- Huson DH, Bryant D. 2006. Application of phylogenetic networks in Evolutionary Studies. *Molecular Biology and Evolution*, 23(2): 254-267.
- Hudson RR, Kreitman M, Aguade M. 1987. A test of neutral molecular evolution based on nucleotide data. *Genetics*, 116: 153-159.
- Jan CC, Fernández-Martínez JM. 2002. Interspecific hybridization, gene transfer and the development of resistance to the broomrape race F in Spain. *Helia*, 25(36): 123-136.
- Joly S, McLenachan PA, Lockhart PJ. 2009. A statistical approach for distinguishing hybridization and incomplete lineage sorting. *American Naturalist*, 174: 54-70.
- Knudsen R, Klemetsen A, Amundsen P-A, Hermansen B. 2006. Incipient speciation through niche expansion: an example from the arctic charr in a subarctic lake. *Proceedings of the Royal Society of Biological Sciences*, 273: 2291-2298.
- Kokkalas S, Xypolias P, Koukouvelas I, Doutsos T. 2006. Postcollisional contractional and extensional deformation in the Aegean region. In: Dilek Y, Pavlides S, eds. *Post Collisional Tectonics & Magmatism in the Mediterranean Region and Asia*. Geological Society of American Sciences, 409, 97-123.
- Koukouvelas I, Kokkalas S, Xypolias P. Surface deformation during the Mw 6.4 (8 June 2008) Movri Mountain earthquake in the Peloponnese, and its implications for the seismotectonics of western Greece. *International Geology Review*, 52(2-3): 249-268.
- Kottelat M, Freyhof J. 2007. *Handbook of the European Freshwater Fishes*. Cornol. Berlin. 2007.
- Kumar S, Stecher G, Tamura, K. 2015. MEGA v.7: Molecular Evolutionary Genetics Analysis Version 7. *Molecular Biology and Evolution* (submitted).
- Leaché AD, McGuire JA. 2006. Phylogenetic relationships of horned lizards (*Phrynosoma*) based on nuclear and mitochondrial data: Evidence for a misleading mitochondrial gene tree. *Molecular Phylogenetics and Evolution*, 39(3): 628-644.
- Librado P, Rozas J. 2009. DnaSP v.5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, 25: 1451-1452.
- Linnen CR, Farrell BD. Mitonuclear discordance is caused by rampant mitochondrial introgression in *Neodiprion* (Hymenoptera: Diprionidae) sawflies. *Evolution*, 61(6): 1417-1438.
- Ludwig A, Congiu L, Pitra C. 2003. Nonconcordant evolutionary history of maternal and paternal lineages in Adriatic sturgeon. *Molecular Ecology*, 12: 3253-3264.
- Machordom A, Berrebi P, Doadrio I. 1990. Spanish barbell hybridization detected using enzymatic markers: *Barbus meridionalis* Risso x *Barbus haasi* Mertens (Osteichthyes, Cyprinidae). *Aquatic Living Resources*, 3(4): 295-303.
- Magalhaes IS, Mwaiko S, Schneider V, Seehausen O. 2009. Divergent selection and phenotypic plasticity during

- incipient speciation in Lake Victoria cichlid fish. *Journal of Evolutionary Biology*, 22: 260-274.
- Marková S, Dufresne F, Manca M, Kotlík P. 2013. Mitochondrial capture misleads about ecological speciation in *Daphnia pulex* complex. *PLoS One*, 8(7): e69497.
- Maddison WP. 1997. Gene trees in species trees. *Systematic Biology*, 46(3): 523-536.
- McKay BD, Zink RM. 2010. The causes of mitochondrial DNA gene tree paraphyly in birds. *Molecular Phylogenetics and Evolution*, 54: 647-650.
- Meng C, Kubatko LS. 2009. Detecting hybrid speciation in the presence of incomplete lineage sorting using gene tree incongruence: A model. *Theory of Population Biology*, 75(1): 35-45.
- Meyer A, Salzburger W, Scharl M. 2006. Hybrid origin of a swordtail species (Teleostei: *Xiphophorus clemenciae*) driven by sexual selection. *Molecular Ecology*, 15: 721-730.
- Nevado B, Koblmüller S, Sturmbauer C, Snoeks J, Usano-Aleman J, Verheyen E. 2009. Complete mitochondrial DNA replacement in a Lake Tanganyika cichlid fish. *Molecular Ecology*, 18: 4240-4255.
- Nosil P, Harmon LJ, Seehausen O. 2009. Ecological explanations for (incomplete) speciation. *Trends in Ecology and Evolution*, 24(3): 145-156.
- Ntageretzi K, Vött A, Fischer P, Hadler H, Emde K, Röloke BR, Willershäuser T. 2015. Paleotsunami history of the Eos Plain (Evrotas River delta, Peloponnese, Greece). *Z. Geomorphol.*, 59(4): 253-273.
- Papanikolaou D, Fountoulis I, Metaxas C. 2007. Active faults, deformation rates and Quaternary paleogeography at Kyparissiakos Gulf (SW Greece) deduced from onshore and offshore data. *Quaternary International*, 172: 14-30.
- Perdices, Economidis PS, Bohlen J, Doadrio I. 2003. Pleistocene effects on the European freshwater fish fauna: double origin of the cobitid genus *Sabanejewia* in the Danube basin (Osteichthyes, Cobitidae). *Molecular Phylogenetics and Evolution*, 26: 289-299.
- Perea S, Böhme M, Zupancic P, Freyhof J, Šanda R, Özulug M, Abdoli A, Doadrio I. 2010. Phylogenetic relationships and biogeographical patterns in Circum-Mediterranean subfamily Leuciscinae (Teleostei, Cyprinidae) inferred from both mitochondrial and nuclear data. *BMC Evolutionary Biology*, 10: 265.
- Rabosky DL, Talaba AL, Donnellan SC, Lovette IJ. 2009. Molecular evidence for hybridization between two Australian desert skinks, *Ctenotus leonhardii* and *Ctenotus quattuordecimlineatus* (Scindidae: Squamata). *Molecular Phylogenetics and Evolution*, 53: 368-377.
- Rambaut A, Drummond AJ. 2007. Tracer v.1.6. Available from <http://beast.bio.ed.ac.uk/Tracer>.
- Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space. *Systematic Biology*, 61(3): 539-542.
- Rohlf FJ, Bookstein FL. 1987. A comment on shearing as a method for "size correction". *Systematic Zoology*, 1987; 36: 357-367.
- Rheindt FE, Edwards SV. 2011. Genetic introgression: an integral but neglected component of speciation in birds. *The Auk*, 128: 620-632.
- Rhymer JM, Simberloff D. 1996. Extinction by hybridization and introgression. *Annual Review of Ecology and Systematics*. 1996; 27: 83-109.
- Rosenberg NA. 2013. Discordance of species trees with their most likely gene trees: a unifying principle. *Molecular Biology and Evolution*. (doi:10.1093/molbev/mst160).
- Sanjur O, Carmona JA, Doadrio I. 2003. Molecular phylogeny of Iberian chub (genus *Squalius*, Cyprinidae) inferred from molecular data. *Molecular Phylogenetics and Evolutions*, 29: 20-30.
- Sinama M, Gilles A, Costedoat C, Corse E, Olivier J-M, Chappaz R, Pech N. 2013. Non-homogeneous combination of two porous genomes induces complex body shape trajectories in cyprinid hybrids. *Frontiers in Zoology*, 10:22.
- Sousa-Santos C, Collares-Pereira MJ, Almada VC. Evidence of extensive mitochondrial introgression with nearly complete substitution of the typical *Squalius pyrenaicus*-like mtDNA of the *Squalius alburnoides* complex

- (Cyprinidae) in an independent Iberian lineage. *Journal of Fish Biology*, 68(SB): 292-301.
- Stadler T. 2009. On incomplete taxon sampling under birth-death models and connection to the sampling-based coalescent. *Journal of Theoretical Biology*, 261: 58-66.
- Stamakis A. 2006. RaxML-VI-HPC: Maximum Likelihood-based phylogenetic Analyses with thousands of taxa and Mixed Models. *Bioinformatics*, 22(21): 2688-2690.
- Stamakis A, Blagojevic F, Nikolopoulos D, Antonopoulos C. 2007. Exploring new search algorithms and hardware for phylogenetics: RaxML Meets the IBM Cell. *The Journal of VLSI Signal Processing*, 48: 271-286.
- Stephens M, Donnelly P. 2003. A comparison of Bayesian methods for haplotype reconstruction from population genotype data. *American Journal of Human Genetics*, 73: 1162-1169.
- Stephens M, Smith N, Donnelly P. 2001. A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics*, 68: 978-989.
- Tang, Q-Y, Liu S-Q, Yu D, Liu H-Z, Danley PD. 2012. Mitochondrial capture and incomplete lineage sorting in the diversification of balitorine loaches (Cypriniformes, Balitoridae) revealed by mitochondrial and nuclear genes. *Zoologica Scripta*, 41(4): 233-247.
- Toews DPL, Brelsford A. 2012. The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology*, 21: 3907-3930.
- Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, 22: 4673-4680.
- Vila M, Vidal-Romaní JR, Björklund M. The importance of time scale and multiple refugia: incipient speciation and admixture of lineages in the butterfly *Erebia triaria* (Nymphalidae). *Molecular Phylogenetics and Evolution*, 36: 249-260.
- Weisrock DWK, Hozak H, Larson A. Phylogeographic analysis of mitochondrial gene flow and introgression in the salamander *Plethodon shermani*. *Molecular Ecology*, 14: 1457-1472.
- Willis SC, Farias IP, Ortí G. 2014. Testing mitochondrial capture and deep coalescence in Amazonian cichlid fishes (Cichlidae: *Cichla*). *Evolution*, 68: 256-268.
- Yu Y, Than C, Degnan JH, Nakleh L. 2011. Coalescent histories on phylogenetic networks and detection of hybridization despite incomplete lineage sorting. *Systematic Biology*, 60(2): 138-149.
- Zink RM, Barraclough G. 2010. Mitochondrial DNA under siege in avian phylogeography. *Molecular Ecology*, 8: 17: 2107-2121.

SUPPLEMENTARY MATERIAL

S1 Table. Additional species used in phylogenetic performance. GenBank accession numbers and labels in phylogenetic trees.

Species	Locality	GenBank accession numbers (labels in phylogenetic trees)
<i>Petroleuciscus ahipsi</i>	Ahipsi River, Russia	MT-CYB: HM560177 (1), HM560178 (2) S7: HM560528 RAG1: HM560455
<i>Petroleuciscus borystheneus</i>	Fotolivos, Greece	MT-CYB: HM560111 S7: HM560499 RAG1: HM560414
<i>Petroleuciscus smyrnaeus</i>	Golcuk Lake, Turkey	MT-CYB: HM560114 S7: HM560501 RAG1: HM560416
<i>Squalius aradensis</i>	Algibre R. Arade Basin, Portugal	MT-CYB: HM560179
<i>Squalius carolitertii</i>	Limia River, Spain	MT-CYB: HM560182 S7: HM560529 RAG1: HM560446
<i>Squalius castellanus</i>	Gallo River, Tajo Basin, Spain	MT-CYB: DQ521423 S7: HM560565 RAG1: HM560575
<i>Squalius malacitanus</i>	Guadiaro River, Spain	MT-CYB: HM560192 S7: HM560531 RAG1: HM560449
<i>Squalius pyrenaicus</i>	Guadalete River, Spain	MT-CYB: HM560198 S7: HM560568 RAG1: HM560579
<i>Squalius valentinus</i>	Magro River, Jucar Basin, Spain	MT-CYB: HM560212 S7: HM560541 RAG1: HM560461
<i>Squalius illyricus</i>	Cetina Basin, Croatia	MT-CYB: HM560183 (1), HM560184 (2) S7: HM560530 RAG1: HM560447
<i>Squalius svallize</i>	Rama Lake, Neretva Basin (1), Zalomka River, (2). Bosnia- Herzegovina	MT-CYB: HM560358 (1), HM560359 (2) S7: HM560538 RAG1: HM560458
<i>Squalius lucumonis</i>	Arno Basin, Italy	MT-CYB: HM560190 RAG1: HM560448
<i>Squalius cephalus</i>	Elbe R. (1), Oder R. (2), Germany	MT-CYB: AJ252807 (1), AY549461 (2)
<i>Squalius laietanus</i>	Matarraña River, Spain	MT-CYB: HM560187, HM560188
<i>Squalius squalus</i>	Vipava R. Isonzo Basin. Slovenia	MT-CYB: HM560204 (1), HM560205 (2)

	(1), Arno River, Italy (2)	S7: HM560537 RAG1: HM560457
<i>Squalius prespensis</i>	Prespa Lake, Greece	MT-CYB: AF090753 (1), HM560197 (2) S7: HM560564 RAG1: HM560457
<i>Squalius orientalis</i>	Uluchaj River, Russia	MT-CYB: HM560195 (1), HM560196 (2) S7: HM560533 RAG1: HM560451
<i>Squalius moreoticus</i>	Stymphalia Lake, Greece	MT-CYB: KY070375-KY070380 (1-6)
<i>Squalius pamvoticus</i>	Arachthos River (1-3), Louros River (4-6), Greece	MT-CYB: KY070381-KY070383 (1-3), KY070384-KY070386 (4-6) S7: KY070513 (1a), KY070514 (1b)

S2 Table. Evolutionary models estimated by jModelTest for the three analysed genes.

GENE	SELECTED MODEL	FREQUENCY OF BASES	RATE MATRIX	GAMMA SHAPE / INVARIABLE SITES
MT-CYB	TIM3+I+G -lnL=5969.3629 k=184	A = 0.2729 C = 0.2970 G = 0.1513 T = 0.2789	R(a) [AC] = 1.9146 R(b) [AG] = 66.5160 R(c) [AT] = 1.0000 R(d) [CG] = 1.9146 R(e) [CT] = 13.8820 R(f) [GT] = 1.0000	Gamma shape = 2.1360 Proportion of invariable sites = 0.5940
S7	TPM3uf+G -lnL=2394.4522 k=116	A = 0.2943 C = 0.1545 G = 0.1982 T = 0.3530	R(a) [AC] = 1.5458 R(b) [AG] = 2.8400 R(c) [AT] = 1.0000 R(d) [CG] = 1.5458 R(e) [CT] = 2.8400 R(f) [GT] = 1.0000	Gamma shape = 0.3030
RAG1	TIM2ef+G -lnL=2816.9850 k=116	Equal	R(a) [AC] = R(b) [AG] = R(c) [AT] = R(d) [CG] = R(e) [CT] = R(f) [GT] =	Gamma shape = 0.0720

S2 Table. Evolutionary models estimated by jModelTest for the three analysed genes.

GENE	SELECTED MODEL	FREQUENCY OF BASES	RATE MATRIX	GAMMA SHAPE / INVARIABLE SITES
<i>MT-CYB</i>	TIM3+I+G	A = 0.2729	R(a) [AC] = 1.9146	Gamma shape = 2.1360
	-lnL=5969.3629	C = 0.2970	R(b) [AG] = 66.5160	Proportion of invariable
	k=184	G = 0.1513	R(c) [AT] = 1.0000	sites = 0.5940
		T = 0.2789	R(d) [CG] = 1.9146	
			R(e) [CT] = 13.8820	
<i>S7</i>	TPM3uf+G	A = 0.2943	R(a) [AC] = 1.5458	Gamma shape = 0.3030
	-lnL=2394.4522	C = 0.1545	R(b) [AG] = 2.8400	
	k=116	G = 0.1982	R(c) [AT] = 1.0000	
		T = 0.3530	R(d) [CG] = 1.5458	
			R(e) [CT] = 2.8400	
<i>RAG1</i>	TIM2ef+G	Equal	R(f) [GT] = 1.0000	
	-lnL=2816.9850		R(a) [AC] =	Gamma shape = 0.0720
	k=116		R(b) [AG] =	
			R(c) [AT] =	
			R(d) [CG] =	
			R(e) [CT] =	
			R(f) [GT] =	

S3 Table. Autapomorphies in the three analysed genes. * transversions.

POSITION <i>MT-CYB</i>																												
POPULATION	18	42	45	60	66	81	88*	93	105	108	117	144	150	151	156	159	162	174	180*	195*	204	228*	231	237	261*	273	279	285
ALFIOS	A	C	T	C	G	C	C	G	G	C	A	G	C	C	G	G	C	C	G	G	C	C	G	T	C	T	C	C
EVROTAS	G	T	C	T	A	T	A	A	C	T	G	A	T	T	A	A	T	T	C	C	T	A	A	C	A	C	T	T
MIRAS	G	T	C	T	A	T	A	A	C	T	A	A	T	T	A	A	T	T	C	C	T	A	A	C	A	C	T	T
PAMISSOS	G	T	C	T	A	T	A	A	C	T	A	A	T	T	A	A	T	T	C	C	T	A	A	C	A	C	T	T

.... Continuation Table S3

POSITION <i>MT-CYB</i>																												
POPULATION	291	303*	309	315	327	351	352	366	372	378	384	390	393	396	405	411	426	432	441	444	462	465	468	474	495	510	513	519
ALFIOS	C	A	C	T	T	A	A	A	A	A	A	T	C	A	A	A	T	G	A	C	T	C	A	C	A	A	T	G
EVROTAS	T	C	T	C	C	G	G	G	G	G	G	C	T	G	G	G	C	A	G	T	C	T	G	C	G	G	C	A
MIRAS	T	C	T	C	C	A	G	G	G	G	G	C	T	G	G	G	C	A	G	T	C	T	G	C	G	G	C	A
PAMISSOS	T	C	T	C	C	A	G	G	G	G	G	C	T	G	G	G	C	A	G	T	C	T	G	T	G	G	C	A

.... Continuation Table S3

POSITION <i>MT-CYB</i>																												
POPULATION	522	552	573	577	579*	580	582	585*	592	598	606	609	615	630*	639	642	651	675*	693*	696	700	702*	708	709	717*	723	724	748
ALFIOS	C	C	T	A	C	A	C	A	C	C	G	G	A	C	C	T	A	A	T	C	A	C	G	T	T	C	T	C
EVROTAS	T	A	C	G	G	G	T	C	T	T	A	A	G	A	T	C	G	C	C	T	G	A	A	C	A	T	C	T
MIRAS	T	A	C	A	G	G	T	C	T	T	A	A	G	A	T	C	G	C	A	T	G	A	A	C	A	T	C	T
PAMISSOS	T	A	C	A	G	G	T	C	T	T	A	A	G	A	T	C	G	C	A	T	G	A	A	C	A	T	C	T

.... Continuation Table S3

POSITION <i>MT-CYB</i> /																									
POPULATION	750	759	774 *	777	784	786	789	792	798	813	816	819	840	849	852	858	861	888	894	897	900	906	915	918	930
ALFIOS	G	G	A	A	C	G	G	T	A	G	G	T	T	T	C	C	A	C	C	A	G	G	A	C	T
EVROTAS	A	A	T	G	T	A	A	G	G	A	A	C	C	C	T	T	G	T	T	G	A	A	C	T	G
MIRAS	A	A	T	G	T	A	A	A	G	A	A	C	C	C	T	T	G	T	T	G	A	A	C	T	G
PAMISSOS	A	A	T	G	T	A	A	A	G	A	A	C	C	C	T	T	G	T	T	G	A	A	C	T	G

.... Continuation Table S3

POSITION <i>MT-CYB</i> /														
POPULATION	1005	1017	1057	1069	1074	1078	1083	1089	1090	1092	1095	1110	1125	1129
ALFIOS	G	G	G	C	T	G	C	C	A	T	T	A	A	C
EVROTAS	A	A	A	T	C	A	A	A	G	C	C	C	G	T
MIRAS	A	A	A	T	C	A	A	A	G	C	C	C	G	T
PAMISSOS	A	A	A	T	C	A	A	A	G	C	C	C	G	T

.... Continuation Table S3

POSITION S7/ POPULATION	68 *	69	72	99	130	266 *	313	414-419	463-477	530	565	581 *	608 *	639
ALFIOS	A	C	C	C	T	T	A	ACGTT	-	G	A	A	A	T
EVROTAS	T	T	T	T	C	A	G	-	TGAGATTAATAATTT	A	G	T	C	C
MIRAS	A	C	C	C	T	T	A	ACGTT	-	G	A	A	A	T
PAMISSOS	A	C	C	C	T	T	A	ACGTT	-	G	A	A	A	T

.... Continuation Table S3

POSITION S7/ POPULATION	642-645	647	653	684 *	732 *	751-752	754	759	802 *	819 *	829	855 *	857	870	876	877-892	945-946
ALFIOS	-	-	T	A	A	TT	T	T	A	G	A	A	G	A	G	-	TG
EVROTAS	TTA	G	-	T	T	-	C	C	C	T	G	T	A	G	A	ACACTATTTTAAGTG	-
MIRAS	-	-	T	A	A	TT	T	T	A	G	A	A	G	A	G	T	TG
PAMISSOS	-	-	T	A	A	TT	T	T	A	G	A	A	G	A	G	T	TG

.... Continuation Table S2

POSITION RAG1/ POPULATION	439	768 *	773	898	948 *	1002	1008	1026	1032	1048 *	1161 *	1169	1173
ALFIOS	G	T	A	T	C	G	C	G	G	C	T	A	G
EVROTAS	A	A	G	C	A	A	T	A	A	A	A	G	A
MIRAS	G	T	A	T	C	G	C	G	G	C	T	A	G
PAMISSOS	G	T	A	T	C	G	C	G	G	C	T	A	G

IV. DISCUSIÓN GENERAL

Entender los procesos que han modelado la distribución y la estructura y diversidad genética actual de los organismos de agua dulce ha sido objeto de interés por parte de los biólogos evolutivos durante las últimas décadas. Los resultados obtenidos en la presente tesis doctoral y desarrollados en los capítulos I a V han contribuido a aumentar el conocimiento sobre algunos de los patrones evolutivos observados en las especies del género *Squalius* Bonaparte, 1837 endémicas de las penínsulas mediterráneas y que pertenecen al gran linaje evolutivo mediterráneo (Sanjur *et al.*, 2003). Estas especies presentan áreas de distribución mucho más restringidas que las especies del linaje Euroasiático (Sanjur *et al.*, 2003), con la presencia de microendemismos, como por ejemplo es el caso de las especies *S. castellanus*, *S. malacitanus* o *S. keadicus*. Por otra parte, en el linaje mediterráneo los procesos evolutivos que presentan son complejos, como consecuencia de factores paleogeográficos y paleoclimáticos intensos y variados, y han dado lugar a patrones filogenéticos y filogeográficos, y a una diversidad que no se encuentra en las especies del linaje Euroasiático.

Las especies que han sido objeto de estudio en la presente tesis constituyen las especies ibéricas del género: *S. castellanus*, *S. carolitertii*, *S. malacitanus*, *S. pyrenaicus* y *S. valentinus*; las especies griegas *S. keadicus* y *S. peloponensis*, perteneciendo esa última especie al linaje Euroasiático de acuerdo a Sanjur *et al.*, (2003) y las especies de la región de Dalmacia, en Bosnia-Herzegovina y Croacia que incluyen *S. illyricus*, *S. microlepis*, *S. svallize*, *S. tenellus* y *S. zrmanjae*. De esta forma, los estudios expuestos a lo largo de los capítulos de la tesis analizan varios aspectos de la historia evolutiva de estas especies de *Squalius*, entre ellos las relaciones filogenéticas y filogeográficas, la descripción de los patrones de estructura y diversidad genética y la presencia de procesos de introgresión genética. Por otro lado, se ha evaluado la respuesta evolutiva de una especie endémica de la península ibérica a la estacionalidad del clima mediterráneo, analizando la influencia de este factor histórico en la estructura y diversidad genética de la misma.

1. Patrones filogenéticos de las especies ibéricas y griegas del género *Squalius*. Evidencia de linajes evolutivos divergentes y capturas mitocondriales antiguas.

Las relaciones filogenéticas de las especies ibéricas del género *Squalius* se han estudiado durante las dos últimas décadas, siendo muchos los trabajos científicos que se han publicado al respecto en base al análisis de marcadores moleculares mitocondriales y nucleares, principalmente alozimas, aunque también secuencias de ADN (Coelho *et al.*, 1995; Alves *et al.*, 1997; Brito *et al.*, 1997; Doadrio y Carmona, 2003, 2004, 2007; Sanjur *et al.*, 2003; Doadrio *et al.*, 2007a,b; Sousa-Santos *et al.*, 2007; Almada y Sousa-Santos, 2010; Waap *et al.*, 2010), así como su marco filogenético dentro de la familia Cyprinidae (Briolay *et al.*, 1998; Zardoya y Doadrio,

1998; Levy *et al.*, 2009; Perea *et al.*, 2010; Geiger *et al.*, 2014). Aun así, muchas de estas relaciones todavía no están totalmente esclarecidas, especialmente a nivel poblacional, como demuestran los resultados obtenidos en los capítulos II y III.

Las relaciones filogenéticas halladas en los capítulos I y II para las especies ibéricas del género *Squalius* son congruentes con trabajos previos a nivel mitocondrial (Sanjur *et al.*, 2003; Doadrio y Carmona, 2003, 2006; Doadrio *et al.*, 2007b; Almada y Sousa-Santos, 2010). Sin embargo, las relaciones filogenéticas con secuencias de genes nucleares aportan nuevos datos y complementan la filogenia obtenida previamente que solamente estudió las relaciones de algunas poblaciones portuguesas y no incluyó las especies endémicas españolas por (Waap *et al.*, 2010).

En esta tesis se ha demostrado que la especie *S. malacitanus* presenta dos linajes genéticos muy bien diferenciados tanto desde el punto de vista mitocondrial como nuclear, uno constituido por las poblaciones mediterráneas (cuencas de los ríos Guadiaro, Guadalmina y Guadaiza) y otros por las poblaciones atlánticas (cuencas de los ríos Barbate, Jara, Vega) y El río Miel, una pequeña cuenca de la bahía de Algeciras. Estos dos grupos divergieron en el límite entre el Plioceno y el Mioceno cuando se produjo la apertura del Estrecho de Gibraltar hace aproximadamente 5.3 Ma. Los apoyos estadísticos de los árboles de genes y árboles de especies realizados en los capítulos I y II fueron muy altos. Las distancias genéticas mitocondriales no corregidas fueron elevadas (4,1-4,5%), siendo incluso mayores que las que presentan otros pares de especies del género reconocidas como válidas (Doadrio y Carmona, 2006; Doadrio *et al.*, 2007a). La presencia de dos linajes tan diferenciados en una región geográfica tan pequeña como es el Estrecho de Gibraltar demuestra la influencia de los factores históricos, paleo-hidrológicos y tectónicos, en los procesos de especiación, y pueden contribuir al mejor conocimiento de una de las regiones más complejas del mundo desde el punto de vista tectónico durante el Cenozoico.

Por otro lado, la especie *S. pyrenaicus* plantea diversas preguntas taxonómicas de interés, como demuestran los resultados expuestos en el Capítulo II y a las cuales se intenta responder en esta Tesis Doctoral. En nuestros análisis hallamos discordancias entre las relaciones filogenéticas inferidas con genes mitocondriales respecto a las relaciones inferidas desde las secuencias de genes nucleares para el conjunto de poblaciones de *Squalius* y especialmente para las diferentes poblaciones que son tradicionalmente asignadas a *S. pyrenaicus* (Sanjur *et al.*, 2003; Doadrio y Carmona, 2006; Doadrio *et al.*, 2007b, Almada y Sousa-Santos, 2010). Los análisis de Bayes y de Máxima Verosimilitud muestran un patrón filogenético incongruente para las topologías mitocondrial y nuclear. De esta forma cuando

utilizamos secuencias de genes mitocondriales las poblaciones de *S. pyrenaicus* presentes en las cuencas del Tajo y del Ebro así como de las pequeñas cuencas próximas a la desembocadura del río Tajo (linaje del norte) forman un grupo monofilético con el resto de poblaciones adscritas a *S. pyrenaicus* (linaje del sur y linaje del río Sado). Sin embargo, con las secuencias de genes nucleares *S. pyrenaicus* no fue un grupo monofilético. Así el linaje del norte de *S. pyrenaicus* estuvo más relacionado con las poblaciones pertenecientes a las especies *S. castellanus* y *S. carolitertii* que con las otras poblaciones de *S. pyrenaicus*. Además, las poblaciones del linaje sur de *S. pyrenaicus* fueron un grupo hermano de las poblaciones de las especies *S. valentinus* y *S. malacitanus* con respecto a la población de *S. pyrenaicus* de la cuenca del río Sado. La topología del árbol concatenando genes nucleares y mitocondriales recuperó una topología similar a la mitocondrial, pero con soportes estadísticos más bajos, que las topologías de los distintos tipos de marcadores analizados de forma independiente. Así, muchas de las relaciones entre las especies ibéricas analizadas que con el análisis bayesiano tuvieron alto o moderadamente alto soporte estadístico fueron recuperadas por los análisis de Máxima Verosimilitud, probablemente debido a las distintas bases metodológicas de ambos métodos filogenéticos. Sin embargo, cuando se utilizaron metodologías de reconstrucción de árboles de especies (*BEAST; Heled y Drummond, 2010), las relaciones filogenéticas entre las especies fueron bien resueltas, cuando se emplearon sólo los seis marcadores nucleares Sin embargo, cuando consideramos todos los genes (mitocondrial y nucleares) los soportes de los nodos fueron bajos para algunas de las relaciones, especialmente las que implicaron a las especies *S. malacitanus*, *S. valentinus* y *S. pyrenaicus*.

Con los genes nucleares todas las especies estuvieron apoyadas estadísticamente por valores de probabilidad posterior altos, excepto el clado que incluye la población del río Zézere y las poblaciones del linaje del norte de *S. pyrenaicus* que tuvo un valor de probabilidad posterior de 0.92. Estas dos poblaciones (Zézere y linaje del norte) incluidas en el árbol de especies son mitocondrialmente diferentes, aunque algunos individuos del río Zézere compartieron ADN mitocondrial con el linaje del norte de *S. pyrenaicus* (Coelho *et al.*, 1995; resultados de esta tesis). Este resultado sugiere que estas poblaciones, Zézere y el linaje del norte de *S. pyrenaicus*, probablemente pertenecen a la misma especie (*S. pyrenaicus*) y que algún patrón evolutivo complejo es el responsable de las discrepancias en las relaciones filogenéticas halladas para este grupo de poblaciones de *Squalius* cuando se emplean diferentes métodos y marcadores.

De esta forma, las discordancias halladas entre los árboles de especies obtenidos a partir de los genes nucleares y a partir de los genes nucleares más el gen mitocondrial revelan lo que podría ser un patrón de hibridación histórica en las especies ibéricas del género *Squalius*.

Considerando que la topología nuclear del árbol de especies refleja de forma más precisa la identidad de las especies, el linaje del sur de *S. pyrenaicus* debería corresponder a un nuevo taxón no descrito, mientras que el linaje del norte sería reconocido como *S. pyrenaicus* de acuerdo a la ubicación de la localidad tipo de la especie (Sintra). Sin embargo, esta asunción no puede ser tomada hasta que se realice un estudio posterior específico para responder esta pregunta taxonómica abierta a partir de los resultados de la presente tesis. Así, en el pasado probablemente se produjo una introgresión mitocondrial desde el linaje sur de “*S. pyrenaicus*” hacia el linaje del norte. Esta introgresión mitocondrial no fue completa en el río Zézere (afluente del tramo bajo de la cuenca del Tajo), probablemente como consecuencia de su especial historia paleogeográfica respecto a otros afluentes de la cuenca. Por tanto, la expansión de los haplotipos mitocondriales de las poblaciones de sur de la Península Ibérica no se extenderían de la misma forma que en otras áreas de la cuenca del río Tajo, y la población del río Zézere conserva en algunos individuos un ADN mitocondrial, distinto probablemente existente previamente a la expansión de las poblaciones del sur de *S. pyrenaicus*. Posteriormente se produciría la divergencia entre los linajes mitocondriales del norte y del sur de “*S. pyrenaicus*”, la cual todavía no es completa ya que aunque en baja frecuencia, ambos linajes todavía comparten haplotipos mitocondriales, como demuestran nuestras redes de haplotipos, siendo el linaje del sur grupo hermano de la especie *S. valentinus* y el linaje del norte de *S. castellanus*.

De esta forma, en esta tesis los métodos bayesianos de reconstrucción de árboles de especies mostraron ser más adecuados para resolver relaciones filogenéticas entre especies que los métodos filogenéticos basados en árboles de genes, como se demuestra para la especie *S. pyrenaicus* en el capítulo II y permitieron descubrir un posible patrón de hibridación pasada que no estaba tan claro en las topologías obtenidas por otras aproximaciones filogenéticas, al no obtener relaciones filogenéticas bien resueltas. Esto pone de manifiesto la influencia de parámetros poblacionales en la propia filogenia de las especies y la robustez de los algoritmos de coalescencia dentro del marco filogenético para resolver las relaciones de parentesco entre las especies. De esta forma, el modelo coalescente multiespecie modela procesos como la separación incompleta de linajes o las coalescencias profundas. Sin embargo, como sugieren los resultados de este mismo capítulo, el método de delimitación de especie bayesiano que se ha utilizado en la presente tesis (*BEAST) no puede incorporar procesos de evolución reticulada o procesos de hibridación (Yu *et al.*, 2012). De esta forma, los valores de probabilidad posterior de los árboles de especies fueron muy bajos cuando estos se reconstruyeron utilizando tanto los marcadores genéticos mitocondriales como nucleares. En el caso de la captura mitocondrial descrita para las especies griegas del género, mostrada en el Capítulo V y que se discute más adelante, el árbol de especies ni siquiera pudo ser construido, ya que las especies implicadas en el proceso de introgresión son muy distantes filogenéticamente.

Los resultados expuestos anteriormente son los más relevantes y novedosos de esta tesis respecto al conocimiento científico actual de las relaciones de las especies del género *Squalius* de la península ibérica. Sin embargo, otros resultados obtenidos en la tesis afianzan resultados previos, como la validez taxonómica de las especies *S. aradensis* (Bogutskaya, Rodrigues y Collares-Pereira, 1988), *S. carolitertii* Doadrio, 1988, *S. castellanus* Doadrio, Perea y Alonso, 2007, *S. malacitanus* Doadrio y Carmona, 2006, *S. pyrenaicus* (Gunther, 1868) y *S. torgalensis* (Bogutskaya, Rodrigues y Collares-Pereira, 1988) en base a su apoyo molecular. Además, también se corrobora la existencia de un linaje evolutivo muy divergente en el río Sado, presentando unas distancias genéticas mitocondriales de 2,2-2,8 % respecto al resto de poblaciones reconocidas como *S. pyrenaicus* (Doadrio y Carmona, 2003; Sanjur *et al.*, 2003; Sousa-Santos *et al.*, 2007; Almada y Sousa-Santos, 2010; Waap *et al.*, 2010). Nuestros resultados también apoyan una divergencia entre el linaje del Mondego respecto al linaje del norte de *Squalius carolitertii* (Doadrio y Carmona, 2003; Santos *et al.*, 2007; Almada y Sousa-Santos, 2010; Waap *et al.*, 2010). Sin embargo, en este caso las distancias genéticas mitocondriales fueron menores del 1%, como se muestra en el capítulo II.

Por su parte, las especies del Mediterráneo oriental del género *Squalius* que pertenecen al linaje evolutivo Mediterráneo están constituidas por tres especies, *S. keadicus* en el Peloponeso griego, y *S. ghigii* y *S. irideus* que habitan en Grecia y Turquía (Kottelat y Freyhof, 2007; Perea *et al.*, 2010; Geiger *et al.*, 2014). La relación filogenética de estas especies con el resto de las especies mediterráneas del género sigue siendo incierta (Perea *et al.*, 2010; Geiger *et al.*, 2014), aunque en esta tesis solamente se ha considerado la posición filogenética de *S. keadicus* en el contexto de una hibridación antigua que dio lugar a una captura mitocondrial en la que además de *S. keadicus* estuvo implicada la especie *S. peloponensis*. Sin embargo, los análisis filogenéticos realizados para las especies adriáticas del género, en los que se han incluido al menos un representante de cada especie perteneciente al linaje Mediterráneo, corrobora la posición filogenética incierta de la especie *S. ghigii*, como se ha observado en previos estudios (Sanjur *et al.*, 2003; Perea *et al.*, 2010; Geiger *et al.*, 2014).

Las especies adriáticas del género *Squalius* que pertenecen al linaje Mediterráneo no presentaron patrones filogenéticos tan complejos como los hallados en la península ibérica, aunque hay que es importante destacar que solamente se analizó un individuo por especie para los genes nucleares (Capítulo III), por lo que no se cuenta con una buena representación de los genotipos nucleares que están presentes en cada una de las especies. Sin embargo, los resultados preliminares obtenidos, en los que se observó una alta congruencia entre las topologías obtenidas por los marcadores mitocondriales y nucleares y unas divergencias entre especies ubicadas en el rango entre el Plioceno Superior y el Pleistoceno Inferior, sugieren la

presencia de procesos de especiación alopátrica reciente dirigidos por la geomorfología del karst Dinárico. Por otro lado, la evaluación de estructura filogeográfica en estas especies, con la presencia de haplotipos mitocondriales compartidos entre todas las poblaciones, indican una alta interconexión entre las poblaciones, especialmente en *S. microlepis*. No obstante, estudios más exhaustivos son necesarios para tratar a fondo estas cuestiones evolutivas.

Respecto al marco filogenético de la especie griega *Squalius keadicus*, perteneciente al linaje Mediterráneo (Sanjur *et al.*, 2003), en base al análisis molecular de un número relativamente elevado de especímenes de varias poblaciones de la especie *S. peloponensis* junto a la única población actual conocida de la especie *S. keadicus* (río Evrotas), corroboramos la existencia de una discordancia mitonuclear que había sido propuesta previamente (Doadrio y Carmona, 1998; Durand *et al.*, 2000). Para ello complementamos los análisis filogenéticos con un análisis morfológico de estas poblaciones mencionadas, como se expone en el Capítulo V (*keadicus*). Así, nuestros resultados muestran que mientras las poblaciones de los ríos Miras y Pamisos están relacionadas mitocondrialmente con la especie *S. keadicus* no ocurre lo mismo nuclearmente, de forma que estas poblaciones se agruparon junto con la población de *S. peloponensis* del río Alfios para dos genes nucleares analizados de forma independiente, *RAG1* y *S7*. Todos los análisis filogenéticos realizados apoyaron con alto soporte estadístico (probabilidades posteriores y bootstrap) las relaciones filogenéticas, tanto mitocondriales como nucleares, expuestas. Los análisis morfológicos llevados a cabo corroboran las relaciones filogenéticas inferidas, especialmente los caracteres merísticos, demostrando la adscripción de las poblaciones de los ríos Miras y Pamisos a la especie *S. peloponensis* y no a *S. keadicus*.

Por este motivo, las relaciones filogenéticas halladas sólo pueden ser fruto de una introgresión mitocondrial desde la especie *S. keadicus* en las poblaciones de los ríos Miras y Pamisos de *S. peloponensis*. Además, todos los individuos de las poblaciones de estos dos ríos presentaron el genoma mitocondrial afín a *S. keadicus*, por lo que fue posible identificar una captura mitocondrial, que también han sido descritas en otros grupos de peces de agua dulce (Freyhof *et al.*, 2005; Nevado *et al.*, 2009; Tang *et al.*, 2012). Sin embargo, el linaje mitocondrial de las poblaciones de los ríos Miras y Pamisos aunque relacionado con el de *S. keadicus* del río EVrotas es diferente y divergieron hace aproximadamente un millón de años de acuerdo a nuestros análisis de datación. Esto da lugar a dos difentes hipótesis, o bien, el genoma introgresado de las poblaciones de los ríos Miras y Pamissos procede de la especie *S. keadicus* y posteriormente ambos linajes mitocondriales divergieron en el Pleistoceno, o bien que la introgresión procede de una especie afín a *S. keadicus* actualmente extinta que compartiría área de distribución con estas dos poblaciones. En el caso de esta última hipótesis la divergencia entre los dos linajes mitocondriales recuperados en los árboles filogenéticos sería más antigua

que el propio evento de introgresión.

La similitud entre las morfologías y el genoma nuclear de las tres poblaciones reconocidas como *S. peloponensis* nos permitió descartar una hipótesis de especiación híbrida entre esta especie y *S. keadicus*. Esto es debido a que las poblaciones híbridas deberían presentar alelos nucleares de ambos parentales y esto no ocurre como demostramos con el análisis filogenético de los genes nucleares con las fases de los alelos resueltas. Sin embargo, incluso las tres poblaciones de *S. peloponensis* analizadas mostraron algunas diferencias morfológicas significativas, como demuestran los análisis morfométricos realizados en el capítulo V. Estas cuatro poblaciones viven actualmente en alopatria y las diferencias morfológicas probablemente entre ellas probablemente aparecieron después del surgimiento de las barreras biogeográficas (sistemas montañosos) que aislaron a las cuatro poblaciones en distintas cuencas hidrográficas del suroeste del Peloponeso a medida que se fueron formando. Hay eso que se llama plasticidad fenotípica para adaptarse a diferentes ambientes.

De esta forma, los datos expuestos en varios capítulos de esta tesis revelan que las introgresiones mitocondriales han sido frecuentes entre las especies del género *Squalius*, como nosotros hallamos en las especies *S. pyrenaicus* y *S. keadicus*, y como se ha descrito para otras especies del género (Durand *et al.*, 2000; Sousa-Santos *et al.*, 2006; Gouksov y Vorburger, 2016). Fenómenos de hibridación intergenérica implicando al género *Squalius* también han sido citadas (Freyhof *et al.*, 2005).

2.- Patrones filogeográficos y estructura poblacional en las especies ibéricas del género *Squalius*

Todas las redes de haplotipos reconstruidas en las distintas especies ibéricas de *Squalius* fueron congruentes con las relaciones filogenéticas inferidas, pero además, pusieron de manifiesto patrones de estructuración genética intraespecífica que no fueron fácilmente observables en los árboles filogenéticos. Estructura genética y geográfica significativa se encontró dentro de las poblaciones mediterráneas de *S. malacitanus* o en las poblaciones del linaje del sur de *S. pyrenaicus*, en los que se observó una clara estructura diferenciando haplogrupos de diferentes cuencas. No fue este el caso para las poblaciones atlánticas de *S. malacitanus*, del linaje norte de *S. carolitertii* o de la especie *S. valentinus*, donde las redes de haplotipos muestran una clara falta de estructura geográfica. En el caso de *S. valentinus*, sin embargo los análisis de AMOVA fueron significativos para una estructura entre las poblaciones, pero esta estructuración fue debida a variación genética entre poblaciones dentro de la misma cuenca hidrográfica y dentro de poblaciones, pero no entre cuencas diferentes.

En dos especies de ibéricas del género *Squalius* se observaron diferentes patrones de estructura genética ante el mismo evento climático, la estacionalidad mediterránea. En la especie *S. valentinus* se aprecia una falta de estructura genética entre las diferentes cuencas hidrográficas donde vive como consecuencia del clima mediterráneo que provoca ciclos de extinciones locales frecuentes y recolonizaciones, que pueden borrar señales históricas de estructura genética como resultado de repetidos cuellos de botella y expansiones poblacionales (Husemann *et al.*, 2012). En la especie *S. pyrenaicus* por el contrario, estos ciclos anuales de cuellos de botella y expansiones poblacionales asociadas al clima mediterráneo han favorecido el mantenimiento de una estructura genética poblacional (metapoblaciones), como también se observa en otras especies de ciprínidos que cohabitan con la especie *S. pyrenaicus* en las cuencas de los ríos Guadiana y Guadalquivir (Sousa-Santos *et al.*, 2014). La posible explicación a este desigual patrón puede ser la diferencia entre la tipología de las cuencas hidrográficas que habitan ambas especies. Así la especie *S. pyrenaicus* habita en las cuencas hidrográficas más grandes de la península ibérica y las pozas estivales pueden pasar periodos de tiempo más largos desconectadas. Por el contrario, los ríos habitados por *S. valentinus* son generalmente muy cortos antes de alcanzar el mar Mediterráneo y muchas veces con pendiente muy marcada, de forma que la mezcla de poblaciones durante los periodos de precipitaciones son más factibles que en el área de distribución de la especie *S. pyrenaicus*. Esta idea es apoyada por el diferente efecto de la estacionalidad mediterránea en las poblaciones de *S. valentinus* dentro de la misma cuenca. Así, las zonas bajas de las grandes cuencas del levante español, especialmente el río Júcar, mostraron una mayor estructuración de sus poblaciones que los tramos medios o altos de las mismas cuencas. Las zonas bajas de estos ríos constituyen llanuras aluviales mientras que los tramos medios presentan un régimen pluvionival y un perfil más abrupto, Esto demuestra la influencia de la geomorfología en la estructura de las poblaciones ibéricas del género *Squalius* así como en las poblaciones de peces agua dulce en general (Carlsson y Nilsson, 2000; Mcglashan y Hughes, 2000; Hughes, 2007; Huey *et al.*, 2008).

Por otro lado, probablemente diferentes procesos demográficos entre *S. valentinus* y *S. pyrenaicus*, o factores intrínsecos a estas especies, sean los responsables de explicar las diferencias encontradas entre los patrones de estructura genética mostrados por ambas especies, ya que está probado el efecto de estos factores sobre la diversidad genética (Frankham, 1996; England *et al.*, 2003; Sousa-Santos *et al.*, 2016).

La estructura genética de las poblaciones de agua dulce también puede verse fuertemente condicionada por la conexión entre poblaciones de distintas cuencas hidrográficas asociadas a procesos climáticos pleistocénicos (glaciaciones). Destacan dos posibilidades de conexiones entre cuencas fluviales, y por tanto dos formas de contacto entre las poblaciones de

peces de agua dulce: conexiones entre tramos bajos de las cuencas o capturas fluviales (Swartz *et al.*, 2007; Costedoat y Gilles, 2009; Unmack *et al.*, 2012). Regresiones marinas y capturas fluviales también han sido postuladas para explicar los patrones filogeográficos actuales del complejo de especies de *S. cephalus* (Costedoat y Gilles, 2009).

Las conexiones entre cuencas hidrográficas a través de los tramos bajos de los ríos normalmente están asociadas a cambios eustáticos en el nivel del mar, estos cambios en la línea de costa pueden producir regresiones marinas ocasionadas por el secuestro del agua en glaciares y capas de hielo aumentando la proporción de tierras emergidas y por tanto la potencialidad de conexión de las desembocaduras de los ríos por sus desembocaduras, especialmente patentes durante el último máximo glacial (Alley *et al.*, 2005; Miller *et al.*, 2005; Sathiamurthy y Voris, 2006). Las reconstrucciones batimétricas realizadas en esta tesis muestran la probabilidad de conexión de las zonas bajas de los ríos del levante español, y por tanto la posible conexión entre las poblaciones de *S. valentinus* (capítulo IV). Sin embargo, esto no ocurre en *S. malacitanus* (capítulo I), debido a la diferencia de anchura de la plataforma continental de las desembocaduras de las cuencas que habitan. La plataforma continental del área del Levante español es ancha y somera (Casas-Sainz y de Vicente, 2009), favoreciendo enormemente las conexiones entre los tramos bajos de las cuencas hidrológicas habitadas por *S. valentinus*. De esta forma, la reconstrucción batimétrica que se presenta en el capítulo IV sugiere que durante el último máximo glacial gran parte de las cuencas más grandes del levante confluyeron en una posible llanura litoral, donde pudieron mezclarse las poblaciones, pero no en la estrecha llanura litoral donde desembocaban las pequeñas cuencas del sur donde actualmente vive *S. valentinus*. Los altos valores de las tasas de migración histórica estimadas para todas las poblaciones de *S. valentinus* junto con la presencia del haplotipo más común y el alelo más común en todas las poblaciones también apoyan la idea de conexiones entre las grandes cuencas del levante español.

Por otro lado, conexiones entre las cabeceras de los ríos a través de capturas fluviales también son factibles, como se ha estimado para *S. pyrenaicus* en el sur de la península ibérica (cuencas de las Cordilleras Béticas) o en la cuenca del Tajo. Las capturas fluviales han facilitado la dispersión de los organismos acuáticos entre cuencas previamente aisladas (Harwood y Hughes, 1998; Craw *et al.*, 2007). Sin embargo los eventos de capturas fluviales no han sido frecuentemente bien documentados, muchas veces porque no se conocen los procesos paleogeográficos que las han originado, por lo que estos eventos geológicos a veces sólo pueden ser inferidos a partir de la estructura genética de las poblaciones (Craw *et al.*, 2007). Nosotros estimamos algunas de estas capturas para las poblaciones de *S. pyrenaicus* del sur de la península ibérica en los capítulos I y II, las cuales se discuten posteriormente.

3.- Patrones de diversidad genética en las especies ibéricas del género *Squalius*

En los estudios llevados a cabo a lo largo de la tesis observamos que en la mayoría de los casos analizados los patrones de diversidad genética que hallamos en las especies ibéricas del género están asociados a factores climáticos. De esta manera, en latitudes más meridionales los patrones de diversidad genética son muy dependientes de la estacionalidad del clima mediterráneo, dando lugar a valores de diversidad más altos respecto a latitudes más septentrionales. De esta forma, nosotros encontramos valores de diversidad genética más altos para *S. pyrenaicus* que en la *S. carolitertii* una especie esta última con una distribución más eptentrional (Sousa-Santos *et al.*, 2016; resultados del capítulo II de esta tesis). Estos valores más elevados de diversidad genética pueden ser debidos a un mayor efecto de la deriva genética y probablemente son muy dependientes de procesos demográficos (Frankham, 1996; Almada y Sousa-Santos, 2010; Sousa-Santos *et al.*, 2014). Por el contrario, en latitudes más septentrionales se observa una reducción de la diversidad genética como se ha mencionado para la especie *S. carolitertii*, probablemente como resultado de fuertes cuellos de botella e incluso extinción local de poblaciones durante las glaciaciones pleistocénicas que ocasionan una drástica pérdida de genotipos (Hewitt, 2004; Provan y Bennet, 2008). Estos patrones latitudinales de diversidad genética a nivel regional en la península ibérica pueden ser extrapolables a nivel de continente europeo, donde las penínsulas mediterráneas muestran valores más altos de diversidad genética respecto a las regiones del centro y norte de la península ibérica (Hewitt, 2004; Schmitt, 2007).

En el levante español se encontraron dos patrones de diversidad genética diferentes para la especie *S. valentinus* en función del marcador genético analizado, mitocondrial o nuclear. Así, se observó un gradiente latitudinal norte-sur de diversidad genética nuclear, desde cuencas de mayor tamaño a cuencas de menor tamaño, que no se reflejó mitocondrialmente. Por el contrario, respecto al gen mitocondrial analizado la diversidad genética fue mayor en las poblaciones de los tramos bajos de las cuencas más grandes y en algunas de las cuencas pequeñas con tramos de agua profundos, principalmente en sus nacimientos, demostrando la influencia de la geomorfología sobre la diversidad genética. Este diferente patrón de diversidad genética entre los dos marcadores responde principalmente a la diferencia en los mismos para reflejar la potencialidad de algunas de las poblaciones de *S. valentinus* para solventar las constricciones impuestas por el clima mediterráneo sobre las poblaciones acuáticas y su efecto en la diversidad genética. En el genoma nuclear es menos marcado que en el mitocondrial entre otras cosas por su mayor tamaño efectivo respecto al genoma mitocondrial, de manera que los alelos tardarán más tiempo en fijarse en las poblaciones (Funk y Omland, 2003). Por el contrario, la diversidad nuclear está reflejando una correlación entre el tamaño de la cuenca hidrográfica y

la diversidad genética, patrón que se ha descrito para otras especies del género *Squalius* (Almada y Sousa-Santos, 2010). Los menores valores de diversidad genética que se encontraron en la región del Levante español para *S. valentinus* se hallaron en las pequeñas cuencas del sur de su área de distribución con un estiaje más marcado, y probablemente sean consecuencia de la mayor vulnerabilidad de estas poblaciones a dicho estiaje, con menor número de refugios disponibles y menor tamaño poblacional, que puede dar lugar a una drástica reducción en el número de individuos, y por tanto a una pérdida de variabilidad genética (Huey *et al.*, 2006; Hughes y Hyllier, 2006).

Pero no en todos los casos los valores de diversidad genética estuvieron determinados por eventos climáticos como principal factor, sino que en el sur de la península ibérica un elemento geomorfológico fue el principal responsable de los mayores valores de diversidad genética hallados en las poblaciones atlánticas de la especie *S. malacitanus* respecto a las mediterráneas. Sin embargo, esta mayor diversidad genética en las poblaciones atlánticas no llevó a una mayor estructura genética con respecto a las poblaciones mediterráneas de *S. malacitanus*. Esto fue consecuencia, como demuestran los análisis de biogeografía realizados de la preponderancia de los eventos de dispersión sobre los de vicarianza en las poblaciones atlánticas de *S. malacitanus*. La alta diversidad genética y falta de estructura de las poblaciones atlánticas de *S. malacitanus* puede explicarse por la presencia de la antigua laguna de la Janda que fue desecada a mediados del siglo XX para pastoreo y agricultura (Recio, 2007). Esta laguna, probablemente actuó como un reservorio de genotipos para las poblaciones de peces del sur de la península ibérica, aumentando por tanto su diversidad genética. Esta laguna endorreica de origen tectónico estuvo alimentada por la cuenca del río Barbate y probablemente permitió la conexión de las poblaciones atlánticas de *S. malacitanus* borrando toda señal de estructura genética entre estas poblaciones (Goy *et al.*, 1995; Luque *et al.*, 1999).

4.- Estimaciones de tiempo de divergencia y biogeografía histórica de las especies de género *Squalius* incluidas en esta tesis

La datación de los tiempos de especiación es especialmente importante dentro de un contexto evolutivo biogeográfico, aunque la obtención de tiempos de divergencia fiables es difícil y lleva asociado fuentes de errores potenciales (Arbogast *et al.*, 2002). La relajación de las premisas metodológicas de las aproximaciones filogenéticas y de reloj molecular ha favorecido la obtención de filogenias y dataciones más precisas, ya que procesos como la estocasticidad subyacente al propio proceso de especiación han podido ser incorporados en los métodos analíticos a través de los árboles de especies y por tanto esto ha permitido obtener hipótesis filogenéticas y estimaciones de tiempos de divergencia más fiables para los procesos de especiación.

Las estimas de tiempos de divergencia que obtenemos en la presente tesis son congruentes con previas estimas obtenidas en otros estudios para el género y sus especies, bien a través de relojes moleculares en función de las distancias genéticas y una tasa de evolución determinada (Zardoya y Doadrio, 1998; Doadrio y Carmona, 2003; Sanjur *et al.*, 2003; Doadrio *et al.*, 2007b), o bien aplicando análisis más específicos de reloj molecular calibrado con fósiles (Levy *et al.*, 2009; Perea *et al.*, 2010). En esta tesis (capítulos I y II) los análisis de *BEAST además de ser utilizados para inferir las relaciones filogenéticas entre las especies también se han usado para estimar los tiempos de divergencia entre las distintas especies ibéricas del género *Squalius*. Esta metodología incorpora en el mismo análisis la reconstrucción filogenética y la datación, permitiendo además trabajar a nivel de especie y a nivel de poblaciones al utilizar algoritmos de coalescencia (Heled y Drummond, 2010). Las estimas que obtuvimos en los capítulos xx y xx fueron congruentes, ya que aunque los valores de las medias de las estimas de los tiempos de divergencia no fueron exactamente iguales, debido a los procesos de estocasticidad de la inferencia bayesiana, sí fueron coincidentes los intervalos de confianza de las mismas. De esta forma, las primeras especies en divergir en el Mioceno Superior - Mioceno Medio fueron *S. aradensis* y *S. torgalensis*. En el Mioceno Medio - Superior se produjo la divergencia de dos linajes de especies bien diferenciados, uno de ellos constituido por las especies *S. carolitertii*, *S. castellanus* y el otro por *S. malacitanus*, *S. pyrenaicus* (Sado y linaje del Sur) y *S. valentinus*. Parece que falta el norte una pequeña explicación de que el norte baila. Las divergencias entre las especies dentro de estos linajes tuvieron lugar en el Plioceno. Estos procesos de divergencia entre las especies ibéricas que han conducido a patrones de especiación alopátrica están fuertemente asociados a los eventos tectónicos y paleogeográficos que dieron lugar a la formación de la actual configuración de la península ibérica y sus cuencas hidrográficas y que han sido descritos en diversos trabajos geomorfológicos (ver por ejemplo Calvo *et al.*, 1993; Vergés and Fernàndez, 2006; De Vicente *et al.*, 2011; Pais *et al.*, 2012; Miller *et al.*, 2013; Torné *et al.*, 2015). Dentro de estos procesos históricos, la separación del margen continental entre África y Europa durante el Mioceno, así como la apertura del Estrecho de Gibraltar en el límite entre el Mioceno y el Plioceno fueron los que parecen haber jugado un papel más importante en modelar la estructura y diversidad genética y los patrones biogeográficos observados en la región occidental del Mediterráneo en el género *Squalius*, al igual que en otros peces de agua dulce (Machordom y Doadrio, 2001; Doadrio y Perdices, 2005).

Aunque el sur de la península ibérica es un área muy compleja desde el punto de vista tectónico (Weijermans, 1991; Krigsman *et al.*, 1999; Miller *et al.*, 2013), los patrones biogeográficos inferidos para las especies que habitan las cuencas próximas al Estrecho de Gibraltar fueron muy claros y presentaron valores de probabilidades marginales relativamente

altos en los análisis de reconstrucción de áreas ancestrales realizados, como se muestra en el capítulo I. La separación del linaje correspondiente a la especie *S. malacitanus* que habita en el sur de la península ibérica tuvo lugar en el límite entre los periodos Serravaliense y Tortonense, hace aproximadamente once millones de años. Divergencias similares han sido estimadas para otros grupos de vertebrados de la región (Fromhage *et al.*, 2004; Carranza *et al.*, 2006). Estas divergencias son más antiguas que la apertura del Estrecho de Gibraltar, por lo que vicarianzas más antiguas que la apertura del estrecho hace 5,33 Ma deben ser propuestas para explicar dichos procesos vicariantes, probablemente asociados con movimientos orogénicos del levantamiento de la Cordillera Bética y reconfiguración de sus cuencas hidrográficas (Betzler *et al.*, 2006; Martin *et al.*, 2009, 2014).

La divergencia de la especie *S. malacitanus* en dos linajes genéticos bien diferenciados uno en las cuencas mediterráneas y otro en las atlánticas coincidiendo con la apertura del Estrecho de Gibraltar (5,33 Ma) supone un patrón biogeográfico similar al que presentan las especies *Aphanius Iberus* y *A. baeticus*, pero diferente al mostrado por el género *Luciobarbus* cuya vicarianza entre las diferentes especies alrededor del Estrecho de Gibraltar tiene una dirección Norte-Sur entre el Norte de África y España y no Este-Oeste entre Mediterráneo y Atlántico español (Perdices *et al.*, 1995; 2001; Machordom y Doadrio, 2001; Doadrio y Perdices, 2005; Almada *et al.*, 2009). Estos patrones vicariantes dando lugar a dos linajes divergentes (atlántico y mediterráneo) en el sur de la península durante el mismo periodo geológico ibérica se han descrito también para otros grupos de vertebrados (Albert *et al.*, 2007).

Sin embargo, *Squalius pyrenaicus* es más reciente en las cuencas del sur de la península ibérica que *S. malacitanus*, ya que su presencia se debe a colonizaciones recientes, como se infirió a través de los análisis de biogeografía histórica realizados en los capítulos I y II. De esta forma, se estimó que la divergencia del linaje del sur de *S. pyrenaicus* con respecto al linaje más cercano tuvo lugar en el Plioceno como consecuencia de la reorganización final de las cuencas hidrológicas ibéricas, tras el paso del endorreísmo miocénico al exorreísmo actual (Calvo *et al.*, 1993; De Vicente *et al.*, 2011; Pais *et al.*, 2012). En nuestros análisis de biogeografía identificamos varios eventos de colonización del sur de la península ibérica desde la cuenca del Guadalquivir a diversas cuencas adyacentes y próximas al Estrecho de Gibraltar (Guadalete, Odiel, Guadalhorca y Guadalfeo) en el Pleistoceno a través de capturas fluviales, como se describe en los capítulos I y II. Colonizaciones desde el Guadiaro a otras cuencas próximas fueron inferidas para las poblaciones mediterráneas de *S. malacitanus*. Conexiones históricas entre los tramos bajos de estas cuencas del sur no fueron posibles, como se deduce de la reconstrucción batimétrica realizada y por la presencia de cañones submarinos profundos (Blanc, 2002; García-Castellanos *et al.*, 2009). Por tanto, las capturas de cabeceras deben ser

supuestas como el medio de conexión más plausible entre estas poblaciones. La proximidad entre las cabeceras de muchas de estas cuencas como por ejemplo los nacimientos de los ríos Guadiaro y Guadaiza debido a la peculiar morfología arqueada de la cuenca del Guadiaro harían factible estas capturas en periodos históricos recientes.

Contactos secundarios pleistocénicos similares a los que se han expuesto para la península ibérica también fueron los posibles responsables de la captura mitocondrial ocurrida en dos poblaciones de la especie *S. peloponensis* en el Peloponeso (Grecia) presentada en el capítulo V. La región del Peloponeso griego es junto con el Estrecho de Gibraltar una de las áreas más complejas de la región Mediterránea; el suroeste del Peloponeso está ubicado cerca de la Trinchera Helénica, una de las zonas más activas tectónicamente al converger las placas Africana y Euroasiática, de forma que presenta numerosas fallas que todavía siguen activas hoy día (Kokkalas *et al.*, 2006; Fountoulis y Mariolakis, 2008; Athanassas y Fountoulis, 2013; Fountoulis *et al.*, 2014). Toda esta complejidad tectónica probablemente promovió contactos secundarios entre poblaciones de agua dulce, resultando en posibles eventos de hibridación. Estos eventos paleogeográficos también serían los responsables de la divergencia pleistocénica de los dos subgrupos monofiléticos hallados en el linaje mitocondrial de *S. keadicus* al configurar la red hidrográfica actual de la región del Peloponeso.

5.- Respuesta evolutiva a la estacionalidad Mediterránea en una especie ibérica del género *Squalius*

No sólo los factores históricos paleogeográficos son los responsables de modelar la historia evolutiva de los organismos vivos, sino que los paleoclimáticos también han tenido mucha influencia en la misma. Uno de estos factores históricos climáticos que han contribuido a modelar la historia evolutiva de las especies ibéricas del género *Squalius* es la fuerte estacionalidad del clima mediterráneo, y es que la mayor parte del territorio de la península se encuentra bajo un régimen climático mediterráneo.

El clima Mediterráneo lleva millones de años establecido en las regiones de tipología mediterránea, entre ellas la península ibérica, pero no es hasta el final del Plioceno, hace aproximadamente 3,2 Ma, cuando los ciclos de sequías en periodos cálidos y lluvias en periodos templados-fríos se establecieron definitivamente (Fauquette *et al.*, 1999; Jiménez-Moreno *et al.*, 2010; Haywood *et al.*, 2016). El origen de la especie *S. valentinus* también data del Plioceno (Doadrio y Carmona, 2006; resultados de esta tesis), de forma que toda su historia evolutiva ha estado ligada al clima mediterráneo, además, esta especie es endémica del Levante español, donde la estacionalidad mediterránea es especialmente acusada (Lázaro *et al.*, 2001). Esto la

convierte en un modelo idóneo para testar su respuesta evolutiva ante la instauración del clima mediterráneo. Para ello se llevaron a cabo aproximaciones de genética del paisaje y modelaje de nicho climático.

Es esperable que el efecto del clima mediterráneo sea más fuerte en ríos más pequeños, especialmente si estos son intermitentes o someros, que en ríos de mayor tamaño o más profundos con más refugios disponibles. La estacionalidad puede exacerbar relaciones densodependientes cuando las poblaciones quedan confinadas en pozas estivales, incrementando así las interacciones bióticas, y afectando por tanto al reclutamiento y a la supervivencia de las poblaciones (Poizat y Crivelli, 1997; Gasith y Ress, 1999; Matthews y Marsh-Matthews, 2003). La estructura y diversidad genética y la historia demográfica puede verse afectada por estas interacciones bióticas que se generan en las pozas estivales. De esta forma, si las poblaciones aisladas por la estacionalidad no sufren un fuerte declive poblacional, que lleve a su extinción, pueden llegar a alcanzar estabilidad demográfica. Este caso es el que se encuentra en las poblaciones de *S. valentinus* de los tramos bajos de algunas de las cuencas hidrográficas que habita, especialmente en la del Júcar, o para algunas poblaciones de los nacimientos de los ríos, tal como se describe en el capítulo IV de la presente tesis. Se ha demostrado que el mantenimiento de estas poblaciones demográficamente estables a lo largo del tiempo ha favorecido el desarrollo de altos niveles de diversidad genética (Sousa-Santos *et al.*, 2014). Por el contrario, las poblaciones de las zonas medias-altas de las cuencas más grandes y aquellas que habitan cuencas pequeñas de regiones más meridionales, pero en las que aparecen manantiales con niveles de agua y temperatura constantes conocidos como ojos o “ullals” presentaron modelos demográficos de expansión poblacional.

Los modelos realizados para analizar cambios en la idoneidad de hábitat a lo largo del tiempo en base a variables bioclimáticas, medida por tanto como nicho climático y asimilable a la distribución potencial de la especie, no real, mostraron pequeños cambios para cada uno de los periodos analizados si observamos las proyecciones al pasado (Último Interglacial, entre 140.000 y 120.000 años; Último Máximo Glacial, hace aproximadamente 21.000 años), al presente (años 1950-2000) y al futuro (año 2080). La distribución más amplia coincide con el Último Interglacial mientras que la más restringida con el Último Máximo Glacial, denotando unas condiciones climáticas más favorables para la especie durante este período y viceversa. Los resultados más interesantes obtenidos de estas proyecciones demuestran que algunas áreas de la distribución de *S. valentinus* presentan valores altos y similares de idoneidad de hábitat para todos los periodos analizados. Estas regiones pueden ser consideradas como áreas de estabilidad para la especie, definidas como zonas que pueden ser habitadas por la especie independientemente del periodo de tiempo (Carnaval *et al.*, 2009). Por lo tanto son capaces de

actuar como refugio y mantener poblaciones viables durante los periodos de sequía estival (Magoulick y Kobza, 2003). Estos refugios son importantes en las regiones de la península ibérica con fuerte estiaje, pero lo son mucho más en regiones áridas o semiáridas donde las sequías pueden prolongarse durante años (Huey *et al.*, 2008; Faulks *et al.*, 2010, 2011). En el Levante español, estas áreas se caracterizan o bien por la presencia de manantiales que llevan agua continuamente (Albufera de Valencia y nacimientos de las pequeñas cuencas meridionales), o bien son áreas que han mantenido un crecimiento demográfico estable a lo largo del tiempo, como los tramos bajos de las cuencas más grandes. A este efecto, los resultados expuestos en el capítulo IV demuestran el valor de los sistemas acuáticos estables, para que las poblaciones de agua dulce que los habitan sean capaces de solventar las severas condiciones impuestas por el clima mediterráneo en esta región (Lázaro *et al.*, 2001). Este modelo es probablemente extrapolable a otras regiones del mundo con fluctuaciones climáticas severas. De hecho, nuestros resultados también demuestran que para la especie *S. valentinus* bajos valores de idoneidad de hábitat favorecen una mayor estructuración de las poblaciones, probablemente a consecuencia de una fragmentación mayor de las mismas.

Las variables bioclimáticas que más contribuyeron para explicar los modelos obtenidos indican la interacción entre la temperatura durante las estaciones del año más cálida y más fría y la precipitación durante los periodos más secos como los factores determinantes de la idoneidad de hábitat para la especie *S. valentinus*. Los análisis de regresión realizados demostraron además que las variables bioclimáticas que tuvieron el mayor efecto sobre la estructura genética fueron las relacionadas con la estacionalidad y condiciones climáticas “extremas” (mayores valores de probabilidad bayesiana marginal). Este hecho vuelve a poner de manifiesto la influencia del clima mediterráneo y sus características sobre la estructura genética de las poblaciones, tal y como se vio de forma empírica para el linaje del sur de *S. pyrenaicus* (Capítulo II) y como ha sido demostrado para otros peces de agua dulce (Magoulick y Kobza, 2003; Sousa-Santos *et al.*, 2014). La asociación de la estacionalidad y las condiciones climáticas “extremas” con una mayor estructuración de las poblaciones sugiere una tendencia a la fragmentación de las mismas en las regiones de tipología climática mediterránea a medida que aumenta la tendencia a la disminución de las poblaciones (Phillandras *et al.*, 2014).

REFERENCES

- Albert EM, Zardoya R, Garcia-Paris M. 2007. Phylogeography and speciation patterns in subterranean worm lizards of the genus *Blanus* (Amphisbaenidae: Blanidae). *Molecular Ecology*, 16(7): 1519-1531.
- Almada VC, Sousa-Santos C. 2010. Comparisons of the genetic structure of *Squalius* populations (Teleostei, Cyprinidae) from rivers with contrasting histories, drainage areas and climatic conditions based on two molecular markers. *Molecular Phylogenetics and Evolution*, 57(2):924-31.
- Alves MJ, Coelho MM, Collares-Pereira MJ. 1997. The *Rutilus alburnoides* complex (Cyprinidae): evidence for hybrid origin. *Journal of Zoological Systematics and Evolutionary Research*, 35:1-10.
- Athanassas C, Fountoulis I. 2013. Quaternary neotectonic configuration of the southwestern Peloponnese, Greece, based on luminescence ages of marine terraces. *Journal of Earth Sciences*, 24(3): 410-427.
- Betzler C, Braga JC, Martin JM, Sánchez-Almazo IM, Lindhorst S. 2006. Closure of a seaway: stratigraphic record and facies (Guadix Basin, Southern Spain). *International Journal of Earth Science*, 95: 431-452.
- Briolay J, Galtier N, Brito RM y Bouvet Y. 1998. Molecular phylogeny of Cyprinidae inferred from cytochrome *b* DNA sequences. *Molecular Phylogenetics and Evolution*, 9: 100-108.
- Brilo RM, Briolay J, Galtier N, Bouvet N, Coelho MM. 1997. Phylogenetic relationships within genus *Leuciscus* (Pisces: Cyprinidae) in Portuguese freshwaters, based on mitochondrial DNA cythocrome *b* sequences. *Molecular Phylogenetics and Evolution*, 8: 435-442.
- Broughton RE, Vedala KC, Crowl TM, Ritterhouse LL. 2011. Current and historical hybridization with differential introgression among three species of cyprinid fishes (genus *Cyprinella*). *Genetica*, 139: 699-707.
- Calvo JP, Daams R, Morales J, López-Martínez N, Agustí J, Anandon P, Armenteros I, Cabrera L, Civis J, Corrochano A, Díaz-Molina M, Elizaga E, Hoyos M, Martín-Suarez E, Martínez J, Moissenet E, Muñoz A, Pérez-García A, Pérez-González A, Portero JM, Robles F, Santisteban C, Torres T, Van der Meulen AJ, Vera JA y Mein P. 1993. Un to date Spanish continental synthesis and paleoclimatic interpretation. *Revista de la Sociedad Geológica de España*, 6: 29-40.
- Carranza R, Harris DJ, Arnold EN, Batista V, González de la Vega JP. 2006. Phylogeography of the lacertid lizard, *Psammodromus algerius*, in Iberia and across the Strait of Gibraltar. *Journal of Biogeography*, 33: 1279-1288.
- Carlsson J, Nilsson J. 2000. Effects of geomorphological structures on genetic differentiation among brown trout populations in a northern boreal river drainage. *Transaction of the American Fisheries Society*, 130(1): 36-45.
- Carnaval AC, Hickerson MJ, Haddad CFB, Rodrigues MT y Moritz C. 2009. Stability predicts genetic diversity in the Brazilian Atlantic Forest hotspot. *Science*, 5915: 785-789.
- Casas-Sainz AM y de Vicente G. 2009. On the tectonic origin of Iberian topography. *Tectonophysics*, 474: 214-235.
- Coelho MM, Brito RM, Brito RM, Pacheco TR, Figueiredo D, Pires A. 1995. Genetic variation and divergence of *Leuciscus pyrenaicus* and *Leuciscus carolitertii* (Pisces, Cyprinidae). *Journal of Fish Biology*, 47: 243-258.
- Costedoat C, Gilles A. 2009. Quaternary pattern of freshwater fishes in Europe: comparative phylogeography and conservation perspective. *The Open Conservation Biology Journal*, 3: 36-48.
- Craw D, Burridge C, Anderson L, waters JM. 2007. Late Quaternary river drainage and fish evolution, Southland, New Zealand. *Geomorphology*, 84(1-2): 98-110.
- Doadrio I, Carmona JA. 1998. Genetic divergence in Greek populations of the genus *Leuciscus* and its evolutionary and biogeographical implications. *Journal of Fish Biology*, 53: 591-613.
- Doadrio I, Carmona JA. 2003. Testing freshwater Lago Mare dispersal theory on the phylogeny relationships of Iberian cyprinid genera, *Chondrostoma* and *Squalius*. *Graellsia*, 59(2-3): 457-473.
- Doadrio I, Carmona JA. 2004. Phylogenetic relationships and biogeography of the genus *Chondrostoma* inferred from mitochondrial DNA sequences. *Molecular Phylogenetics and Evolution*, 33: 802-815.

- Doadrio I, Carmona JA. 2006. Phylogenetic overview of the genus *Squalius* (Actinopterygii, Cyprinidae) in the Iberian Peninsula, with description of two new species. *Cybum*, 30(3): 199-214.
- Doadrio I, Kottelat M, Sostoa A. 2007a. *Squalius laietanus*, a new species of cyprinid fish from north-eastern Spain and southern France (Teleostei, Cyprinidae). *Ichthyological Exploration of Freshwaters*, 18(3):247-256.
- Doadrio I, Perdices A. 2005. Phylogenetic relationships among the Ibero-African cobitids (*Cobitis*, Cobitidae) based on cytochrome b sequence data. *Molecular Phylogenetics and Evolution*, 37(2): 484-493.
- Doadrio I, Perea S, Alonso F. 2007b. A new species of the genus *Squalius* Bonaparte, 1837 (Actinopterygii, Cyprinidae) from the Tagus River Basin (Central Spain). *Graellsia*, 63(1): 89-100.
- Durand JD, Unlü E, Doadrio I, Pipoyan S, Templeton AR. 2000. Origin, radiation, dispersion and allopatric hybridization in the chub *Leuciscus cephalus*. *Proceedings of the Royal Society of London*, B267: 1687-1697.
- England PR, Osler GHR, Woodworth LM, Montgomery ME, Briscoe DA, Frankham R. 2003. Effects of intense versus diffuse population bottlenecks on microsatellite genetic and evolutionary potential. *Conservation Genetics*, 4(5): 595-604.
- Faulks LK, Gilligan DM y Beheregay LB. 2010. The role of anthropogenic vs natural in-stream structure in determining connectivity and genetic diversity in an endangered freshwater fish, Macquarie perch (*Macquaria australasica*). *Evolutionary applications*, 4: 589-601.
- Fauquette S, Suc J-P, Guiot J, Diniz F, Feddi N, Zheng Z, Bessais E y Drivaliari A. 1999. Climate and biomes in the west Mediterranean area during the Pliocene. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 152:15-36.
- Fountoulis I, Mariolakos I. 2008. Neotectonic folds in the central-western Peloponnese, Greece. *Z. Dtsch. Ges. Geowiss*, 159(3): 485-494.
- Fountoulis I, Mariolakos I, Ladas I. 2014. Quaternary basin sedimentation and geodynamics in SW Peloponnese (Greece) and late stage uplift of Taygetos Mt. *B. Geofis. Teor. Appl.*, 55(2): 303-324.
- Frankham R. 1996. Conservation genetics. *Annual Review of Genetics*, 29: 305-327.
- Fromhage L, Vences M, Veith M. 2004. Testing alternative vicariance scenarios in Western Mediterranean discoglossid frogs. *Molecular Phylogenetics and Evolution*, 31: 308-322.
- Funk DJ, Omland KE. 2003. Species-level paraphyly and polyphyly: Frequency, causes, and consequences, with insights from Animal Mitochondrial DNA. *Annual Review of Ecology and Evolutionary Systematics*, 34: 397-423.
- Freyhof J, Dietmar L, Pitra C, Ludwig A. 2005. Molecules and morphology: Evidence for introgression of mitochondrial DNA in Dalmatian cyprinids. *Molecular Phylogenetics and Evolution*, 37: 347-354.
- García-Castellanos D, Estrada F, Jiménez-Munt I, Gorini C, Fernández M, Vergés J, De Vicente R. 2009. Catastrophic flood of the Mediterranean after the Messinian Salinity Crisis. *Nature*, 462(10): 778-782.
- Gasith A y Resh VH. 1999. Streams in Mediterranean climate regions: abiotic influences and biotic responses to predictable seasonal events. *Annual Review of Ecology and Systematics*, 30: 51-81.
- Geiger MF, Herder F, Monaghan M, Almada V, Barbieri R, Bariche M, Berrebi P, Bohlen J, Casal-Lopez M, Delmastro G, Denys G, Dettai A, Doadrio I, Kalogianni E, Kärst H, Kottelat M, Kovacic M, Laporte M, Lorenzoni M, Marcic Z, Özuluğ M, Perdices A, Perea S, Persat H, Porcellotti S, Puzzi C, Robalo J, Sanda R, Schneider M, Slechtova V, Stumboudi M, Walter S, Freyhof J. 2014. Spatial heterogeneity in the Mediterranean Biodiversity Hotspot affects barcoding accuracy of its freshwater fishes. *Molecular Ecology Resources*, 14: 1210-1221.
- Gouksov A, Vorburger C. 2016. Postglacial recolonizations, watershed crossing and human translocations shape the distribution of chub lineages around the Swiss Alps. *BMC Evolutionary Biology*, 16(1): 185.
- Haywood AM, Dowsett HJ, Dolan AM. 2016. Integrating geological archives and climate models for the mid-Pliocene warm period. *Nature Communications*, 7:10646.
- Heled J, Drummond A. 2010. Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution*, 27(3): 570-580.
- Huey JA, Hughes JM y Baker AM. 2006. Patterns of gene flow in two species of eel-tailed catfish, *Neosilurus hyrtii*

- and *Porochilus argenteus* (Siluriformes: Plotosidae), in western Queensland's dryland rivers. *Biological Journal of the Linnean Society*, 87: 457-467.
- Hughes JM, Hyllier MJ. 2006. Mitochondrial DNA and allozymes reveal high dispersal abilities and historical movement across drainage boundaries in two species of freshwater fishes from inland rivers in Queensland, Australia. *Journal of Fish Biology*, 68(B): 270-291.
- Hughes PD, Woodward JC. 2008. Timing of glaciation in the Mediterranean mountains during the last cold stage. *Journal of Quaternary Science*, 23: 575-588.
- Hewitt GM. 2004. Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transaction of Royal Society Series B, Biological Sciences*, 359: 183-195.
- Husemann M, Ray JW, King RS, Hooser EA, Danley PD. 2012. Comparative biogeography reveals differences in population genetic structure of five species of stream fishes. *Biological Journal of the Linnean Society*, 107: 867-885.
- Jiménez-Moreno G, Suc J-P y Fauquette S. 2010. Miocene to Pliocene vegetation reconstruction and climate estimates from the Iberian Peninsula from pollen data. *Review of Paleobotany and Palynology*, 162: 403-415.
- Kokkalas S, Xypolias P, Koukouvelas I, Doutsos T. 2006. Postcollisional contractional and extensional deformation in the Aegean region. In: Dilek Y, Pavlides S, eds. *Post Collisional Tectonics & Magmatism in the Mediterranean Region and Asia*. Geological Society of American Sciences, 409, 97-123.
- Kottelat M, Freyhof J. 2007a. Handbook of the European Freshwater Fishes. Cornol. Switzerland. 646 pp.
- Krijgsman W, Hilgen FJ, Raffi I, Sierro FJ, Wilson DS. 1999. Chronology, causes and progression of the Messinian Salinity Crisis. *Nature*, 400: 652-655.
- Lázaro R, Rodrigo FS, Gutiérrez L, Domingo F y Puigdefábregas J. 2001. Analysis of a thirty year rainfall record (1967-1997) from semi-arid SE Spain: a plant ecological perspective. *Journal of Arid Environments*, 48: 373-395.
- Levy A, Doadrio I, Almada VC. 2009. Historical biogeography of European leuciscins (Cyprinidae): evaluating the Lago Mare dispersal hypothesis. *Journal of Biogeography*, 36(1): 55-65.
- Machordom A, Doadrio I. 2001b. Evidence of a Cenozoic Betic-Kabilian connection based on freshwater fish phylogeography (*Luciobarbus*, Cyprinidae). *Molecular Phylogenetics and Evolution*, 18(2): 252-263.
- Mcglashan DJ, Hughes JM. 2000. Reconciling patterns of genetic variation with stream structure, earth history and biology in the Australian freshwater fish *Craterocephalus stercusmuscarum* (Atherinidae). *Molecular Ecology*, 9(11): 1737-1751.
- Magoulick DD, Kobza RM. 2003. The role of refugia for fishes during drought: a review and synthesis. *Freshwater Biology*, 48: 1186-1198.
- Martín JM, Braga JC, Aguirre J, Puga-Bernabéu A. 2009. History and Evolution of the North-Betic Strait (Prebetic Zone, Betic Cordillera): a narrow, early Tortonian, tidal-dominated, Atlantic-Mediterranean marine passage. *Sedimentary Geology*, 216: 80-90.
- Martín JM, Puga-Bernabéu A, Aguirre J, Braga JC. 2014. Miocene Atlantic-Mediterranean seaways in the Betic cordillera (southern Spain). *Revista de la Real Sociedad Geológica de España*, 27(1): 175-186.
- Matthews WJ, Marsh-Matthews E. 2003. Effects of drought on fish across axes of space, time and ecological complexity. *Freshwater Biology*, 48: 1232-1253.
- Miller KG, Kominz MA, Browning JV, Wright JD, Mountain GS, Katz ME, Sugarman PJ, Cramer BS, Christie-Blick N, Pekar SF. 2005. The Phanerozoic record of global sea-level change. *Science*, 310: 1293-1298.
- Miller MS, Allam AA, Becker TW, Di Leo JF, Wookey J. 2013. Constraints on the tectonic evolution of the westernmost Mediterranean and northwestern Africa from shear wave splitting analysis. *Earth and Planetary Science Letters*, 375(1): 234-243.
- Nevado B, Koblmüller S, Sturmbauer C, Snoeks J, Usano-Aleman J, Verheyen E. 2009. Complete mitochondrial DNA replacement in a Lake Tanganyika cichlid fish. *Molecular Ecology*, 18: 4240-4255.

- Pais JA, Cunha PP, Pereira D, Legoinha P, Dias R, Moura D, Brum da Silveira A, Kullberg JC, González-Delgado JA. 2012. The Paleogene and Neogene of western Iberia (Portugal). *A Cenozoic record in the European Atlantic domain. Springer Briefs in Earth Sciences*. Springer. Berlin.
- Perdices A, Carmona JA, Fernández-Delgado C, Doadrio I. 2001. Nuclear and mitochondrial data reveal high genetic divergence among Atlantic and Mediterranean populations of the Iberian killifish *Aphanius iberus* (Teleostei: Cyprinodontidae). *Heredity*, 87(3): 314-324.
- Perdices A, Machordom A, Doadrio I. 1995. Allozyme variation of African and Iberian populations of the genus *Cobitis* Linnaeus, 1758 (Osteichthyes, Cobitidae). *Journal of Fish Biology*, 47: 707-718.
- Perea S, Böhme M, Zupancic P, Freyhof J, Šanda R, Özulug M, Abdoli A, Doadrio I. 2010. Phylogenetic relationships and biogeographical patterns in Circum-Mediterranean subfamily Leuciscinae (Teleostei, Cyprinidae) inferred from both mitochondrial and nuclear data. *BMC Evolutionary Biology*, 10: 265.
- Phillandras CM, Nastos PT, Kapsomenakis J, Douvis KC, Tselioudis G y Zerefos CS. 2011. Long term precipitation trends and variability within the Mediterranean region. *Natural Hazards and Earth Systems Sciences*, 11: 3235-3250.
- Poizat G y Crivelli AJ. 1997. Use of seasonally flooded marshes by fish in a Mediterranean wetland: timing and demographic consequences. *Journal of Fish Biology*, 51: 106-119.
- Provan J, Bennett KD. 2008. Phylogeographic insights into cryptic glacial refugia. *Trends in Ecology and Evolution*, 23(10): 564-571
- Sanjur OI, Carmona JA, Doadrio I. 2003. Evolutionary and biogeographical patterns within Iberian populations of the genus *Squalius* inferred from molecular data. *Molecular Phylogenetics and Evolution*, 29(1): 20-30.
- Schmitt T, Hewitt G. 2004. The genetic pattern of populations threat and loss: a case study of butterflies. *Molecular Ecology*, 13: 21-31.
- Sousa-Santos C, Collares-Pereira MJ, Almada VC. 2007. Evidence of extensive mitochondrial introgression with nearly complete substitution of the typical *Squalius pyrenaicus*-like mtDNA of the *Squalius alburnoides* complex (Cyprinidae) in an independent Iberian drainage. *Journal of Fish Biology*, 68(SB): 292-301.
- Sousa-Santos, C., Robalo, J.I., Pereira, A.M., Branco, P., Santos, J.M., Ferreira, M.T., Sousa, M, Doadrio, I. 2016. Broad-scale sampling of primary freshwater fish populations reveals the role of intrinsic traits, inter-basin connectivity, drainage area and latitude on shaping contemporary patterns of genetic diversity. *PeerJ*. DOI 10.7717/peerj.1694.
- Swartz ER, Skelton, PH y Bloomer P. 2007. Sea-level changes, river capture and the evolution of populations of the Eastern Cape and fiery redbins (*Pseudobarbus after* and *Pseudobarbus phlegethon*, Cyprinidae) across multiple river systems in South Africa. *Journal of Biogeography*, 34: 2086-2099.
- Tang, Q-Y, Liu S-Q, Yu D, Liu H-Z, Danley PD. 2012. Mitochondrial capture and incomplete lineage sorting in the diversification of balitorine loaches (Cypriniformes, Balitoridae) revealed by mitochondrial and nuclear genes. *Zoologica Scripta*, 41(4): 233-247.
- Unmack PJ, Bagley JC, Adams M, Hammer MP y Johnson JB. 2012. Molecular phylogeny and phylogeography of the Australian freshwater fish genus *Galaxiella*, with an emphasis on dwarf *Galaxias* (*G. pusilla*). *PloS One*, 7: e38433.
- Vergés, J., Fernández, M. (2006) Ranges and basins in the Iberian Peninsula: their contribution to the present topography. *European lithosphere dynamics: Geological Society, London. Memoirs* (ed. by Gee, G.D. & Stephenson, R.A.). pp 223-234.
- Waap S, Amaral AR, Gomes B, Coelho MM. 2010. Multilocus species tree of the chub genus *Squalius* (Leuciscinae: Cyprinidae) from western Iberia: new insights into its evolutionary history. *Genetica*, 139: 1009-1018.
- Weijermans R. 1991. Geology and tectonics of the Betic Zone, southeast Spain. *Earth Science Review*, 31: 153-236.
- Yu Y, Degnan JH, Nakhleh L. 2012. The probability of a gene tree topology within a phylogenetic network

applications to hybridization detection. PLoS Genetics, 8(4): e1002660.

Zardoya R y Doadrio I. 1998. Phylogenetic relationships of Iberian Cyprinidae: Systematic and biogeographical implications. *Proceedings of the Royal Society of London Series B.*, 265: 1365-1372.

V. CONCLUSIONES

1. Nuestros resultados demuestran una fuerte influencia de los factores tectónicos y climáticos sobre la historia evolutiva de las especies mediterráneas del género *Squalius* que se refleja en los patrones filogenéticos y filogeográficos hallados.
2. La taxonomía de las especies ibéricas del género *Squalius* sigue siendo una cuestión pendiente por resolver como sugiere la presencia de linajes genéticos muy divergentes en varias especies, entre ellas *S. pyrenaicus* y *S. malacitanus*.
3. La presencia de discordancias entre las topologías mitocondrial y nuclear en la especie ibérica *S. pyrenaicus* parece responder a un patrón de hibridación antigua que dio lugar a una introgresión mitocondrial entre los linajes del norte y del sur de esta especie seguida de una posterior divergencia entre estos dos linajes desde el punto de vista mitocondrial. Sin embargo, las relaciones filogenéticas nucleares halladas parecen indicar que estos dos linajes probablemente constituyan dos taxones distintos.
4. La especie *S. malacitanus* también presenta dos linajes evolutivos muy divergentes tanto mitocondrial como nuclearmente (atlántico y mediterráneo) que se separaron hace aproximadamente cinco millones de años.
5. Se encontraron dos patrones evolutivos diferentes en dos especies del género *Squalius* en el sur de la península ibérica, diferenciados principalmente por el tiempo de colonización de esta región por ambas especies. La especie *Squalius malacitanus* es más antigua en la región y su historia evolutiva se vio principalmente afectada por dos eventos tectónicos: el levantamiento de la cordillera Bética, que lo aisló del resto de la península ibérica, y la apertura del Estrecho de Gibraltar, que dio lugar a la divergencia entre las poblaciones atlánticas y mediterráneas.
6. La colonización del sur de la península ibérica por *S. pyrenaicus* es más reciente y consecuencia de múltiples colonizaciones desde la cuenca del Guadalquivir, posiblemente vía capturas fluviales, ya que conexiones entre los tramos bajos de los ríos no parecen plausibles de acuerdo a la reconstrucción batimétrica realizada.
7. Además de los eventos pleistocénicos de colonización desde el Guadalquivir a otras cuencas del sur de la península, los análisis de reconstrucción de áreas ancestrales infirieron para el linaje del sur de *S. pyrenaicus* una primera vicarianza entre las cuencas del Guadiana-Gilão y la cuenca del Guadalquivir a medida que ambas cuencas se fueron formando.

8. En el linaje del norte de *S. pyrenaicus* se identificaron varios eventos de vicarianza y dispersión entre la cuenca del Tajo, la cuenca del Lis y otras pequeñas cuencas próximas a su desembocadura durante el Pleistoceno.
9. En la especie *Squalius carolitertii* se identificó un patrón de vicarianza entre los linajes genéticos del norte y del Mondego. Además, se infirieron colonizaciones pleistocénicas desde el Duero, que probablemente actuó como refugio glacial, hacia otras cuencas del cuadrante noroccidental de la península ibérica.
10. La especie *Squalius valentinus* estuvo filogenéticamente emparentada con el linaje del sur de *S. pyrenaicus*, y la divergencia entre ambas especies se explica por vicarianzas entre las cuencas hidrográficas que habitan.
11. Los patrones de estructura genética difirieron entre las distintas especies ibéricas de *Squalius* analizadas. *S. pyrenaicus* y *S. malacitanus* presentaron estructura genética significativa asociada a la presencia de distintos linajes evolutivos dentro de las especies. A su vez, solo los dos linajes evolutivos de *S. pyrenaicus* y el linaje mediterráneo de *S. malacitanus* mostraron estructura genética, no así los linajes de *S. carolitertii* o el linaje atlántico de *S. malacitanus*.
12. La estructura de los linajes de *S. pyrenaicus*, al igual que sus altos valores de diversidad genética, respecto a la especie *S. carolitertii*, probablemente se debe a la presencia de una estructura metapoblacional asociada a la presencia continuada de fragmentación de poblaciones y cuellos de botella anuales, derivando en la fijación de distintos genotipos en las distintas poblaciones aisladas.
13. En *S. carolitertii* la falta de estructura y la baja diversidad genética puede asociarse a las glaciaciones pleistocénicas, que probablemente dieron lugar a fuertes reducciones poblacionales y extinción de poblaciones.
14. A pesar de habitar en un régimen climático mediterráneo que favorece la fragmentación de poblaciones, en *S. valentinus* la falta de estructura genética entre cuencas, que sí fue significativa entre poblaciones dentro de cuencas, se debe probablemente a la alta conectividad entre poblaciones cuando se restauran las precipitaciones tras un periodo de sequía. Por otro lado, conexiones históricas pleistocénicas por los tramos bajos de los ríos, a consecuencia de cambios en el nivel del mar también han enmascarado la estructura genética asociada a la estacionalidad mediterránea en esta especie.

15. En *S. malacitanus*, la laguna de la Janda jugó un papel clave para explicar la alta diversidad genética hallada en las poblaciones atlánticas de esta especie en comparación a las poblaciones mediterráneas. Esta laguna probablemente supuso un importante reservorio de diversidad genética para las poblaciones de peces de agua dulce en general del sur de la península ibérica.

16. Nuestros resultados evidencian el importante papel de condiciones ambientales estables en refugios de hábitat provistos por pequeños manantiales kársticos para mantener poblaciones con una alta variabilidad genética, como ocurre en la especie *S. valentinus* en el levante español.

17. Las fluctuaciones de las condiciones climáticas ocurridas durante el Pleistoceno han favorecido el desarrollo de dos modelos demográficos diferentes para *S. valentinus*. Las poblaciones de pequeños ríos con estacionalidad mediterránea extrema son más susceptibles a los ciclos de sequía/lluvia y sufren fuertes reducciones poblacionales. Las poblaciones de hábitats que llevan permanentemente agua, como en los ríos más grandes o en los pequeños manantiales kársticos, mostraron estabilidad demográfica o expansión poblacional.

18. Las variables bioclimáticas que presentaron mayor efecto sobre la idoneidad de hábitat a lo largo de todos los periodos históricos analizados fueron la presencia de precipitaciones durante el periodo más cálido y un estrecho rango de temperatura anual a lo largo de todos los periodos analizados. Las variables bioclimáticas relacionadas con la estacionalidad mediterránea favorecen una fragmentación de las poblaciones.

19. Las especies adriáticas del género *Squalius* mostraron patrones de especiación alopátrica asociada a las etapas finales de formación del karst Dinárico en el Plioceno Superior y Pleistoceno Inferior.

20. La estructura filogeográfica de especies como *S. microlepis*, *S. illyricus* o *S. svallize* probablemente está influenciada por la compleja configuración del karst Dinárico como consecuencia de conexiones subterráneas que conectan cuencas aparentemente aisladas en superficie.

21. Los resultados de esta tesis corroboran la existencia de una captura mitocondrial a través de análisis morfológicos y moleculares entre dos especies griegas del género *Squalius* probablemente debida a contactos secundarios entre cuencas durante el final del Plioceno y Pleistoceno, ya que la región del Peloponeso es una zona muy compleja desde el punto de vista tectónico.

22. El genoma mitocondrial de las poblaciones introgresadas constituye un grupo recíprocamente monofilético respecto a la especie parental *S. keadicus*. Esto podría sugerir o bien que los dos genomas han divergido completamente o que una especie filogenéticamente afín a *S. keadicus* pero extinta sería la especie parental implicada en el proceso de hibridación.

23. Los distintos procesos de hibridación inferidos en esta tesis sugieren que las barreras de aislamiento reproductor en el género *Squalius* son difusas.